

CEMP-RT

DEPARTMENT OF THE ARMY
U.S. Army Corps of Engineers
Washington, DC 20314-1000

ETL 1110-1-175

Technical Letter
No. 1110-1-175

30 June 1997

Engineering and Design
**PRACTICAL ASPECTS OF APPLYING GEOSTATISTICS
AT HAZARDOUS, TOXIC, AND RADIOACTIVE WASTE SITES**

1. Purpose. The principal purpose of this ETL is to introduce the reader to geostatistical techniques and to demonstrate their basic utility with respect to HTRW site investigations. The ETL also will include a discussion of statistical concepts that support the science of geostatistics. Practical aspects of geostatistical techniques will be discussed in two ways. First, practical references will be made, when appropriate, during the discussion of statistical concepts, and second, examples describing several aspects of the use of geostatistical techniques in HTRW site investigations will be presented and discussed in a section of this ETL specifically dedicated to providing working examples. This ETL also will include a brief literature and software review; review of geostatistical applications; comparison of information that is generated with geostatistical methods to that information obtained using classical statistical methods; and some more recent geostatistical methods, such as conditional simulation.

2. Applicability. This letter applies to all USACE commands having HTRW investigation, design, and remedial action responsibility within the military or civil works programs.

3. References. Documents referenced in this ETL are listed. Appendix A contains additional references useful in geostatistical application.

a. EM 200-1-2, Technical Project Planning Guidance for HTRW Data Quality Design.

b. ASTM D-5922, Standard Guide for Analysis of Spatial Variation in Geostatistical Site Investigations.

c. ASTM D-5549, Standard Guide for Content of Geostatistical Site Investigations.

4. Distribution Statement. Approved for public release, distribution is unlimited.

5. Discussion.

a. Geostatistics is a powerful tool to assess relationships among data obtained from various locations. It allows optimization of sample spacing and frequency. More importantly, geostatistics also allows one to effectively estimate parameter values in areas between actual sample points and quantify the uncertainty of the estimated values. This can be very valuable in risk management and design decision making. This ETL builds upon the principles introduced in EM 200-1-2.

b. The ETL contains examples which illustrate the statistical principles discussed throughout the document. Not every application of geostatistics to HTW projects could be illustrated, however, and the user must be aware of the basic principles and seek appropriate applications. Specific examples of typical cost-effective applications of geostatistics are also given here.

(1) Geostatistics, by the construction of a variogram based on preliminary sampling, can be used to determine the **typical separation of sampling points that delineate uncorrelated data**. The range of the variogram is used as a basis for selecting a sample spacing that minimizes costs and provides independent data for determining, for example, average exposure values for risk assessment. First, an adequate number of preliminary samples are analyzed from the site (refer to section 4-3). Second, a variogram is constructed using techniques described in Chapter 4. Third, the range of the variogram, as defined in section 2-3 is determined. Lastly, the range or some multiple or fraction of it, is chosen for future sample spacing. The variogram should be updated as new data are collected. For example, the variogram may indicate data spaced more than 200 ft apart are uncorrelated. Closure sampling may then be proposed to be spaced every 200 ft or more along an excavation. Smaller spacing results in unnecessary duplication of information and unneeded expenditure of funds.

(2) Geostatistics, through block kriging, can yield **estimates of the average concentrations to be encountered** in a typical daily excavation area/volume. For applications such as excavation of near surface contamination, two-dimensional block kriging could be used to estimate mean contaminant concentration for specific excavation areas. Although this document does not address three-dimensional block kriging for estimating mean concentrations within given volumes, additional guidance and tools for three-dimensional kriging are available through references cited in Appendix A. Alternatively, one can use two-dimensional block kriging to estimate mean concentrations in different layers within a given volume. These estimates can then be averaged to approximate the overall average concentration within the entire volume. This assumes adequate data exist to perform the two-dimensional block kriging at the different depths. To perform two-dimensional block kriging, adequate site characterization data are collected (refer to section 4-4). Second, the data gathered from the areas of interest are used to construct a variogram, as described Chapter 4. Third, the variogram is

modeled as described in section 4-6. Lastly, the model is used to perform block kriging, as described in section 2-4 for blocks of a size comparable to the daily excavation area/volume. The block-kriged values can then be used for estimating the treatment plant loading, etc., related to that block. The kriging also quantifies the possible variance in the average concentration for each block that can be used to manage the risk of operating a treatment plant.

(3) Exposure concentrations for risk assessment purposes can be computed, using geostatistics, even though the site characterization data are somewhat clustered or were collected using biased sampling strategies. Assuming the data are already available and adequate in number (refer to section 4-4), the first step is to compute a sample variogram, as described in Chapter 4. Second, the variogram is modeled as described in section 4-6. Next, this model is used in performing a block kriging operation over the inferred exposure area, as described in section 2-3. Finally, the block kriging value can be used, along with the kriging variance, **to determine the exposure point concentration**, assuming the data were normally distributed (or were transformed to be normally distributed).

(4) The last example describes the use of geostatistics **to quantify project risk for excavation or treatment volumes**. Even with ample site characterization point data (borings or wells), the limits of the treatment zone are imperfectly defined. Geostatistics allows one to evaluate the risk that the size, and therefore cost, of the remediation may be larger or smaller than expected. First, site characterization is performed and adequate data are collected (as described in section 4-4). Second, the data are transformed by assigning a value of one or zero, depending on whether the value is above or below, respectively, a given clean-up value or other criteria. Third, the transformed data are then used to construct a variogram as described in Chapter 4. Fourth, this variogram is modeled as described in section 4-6. Next, this model is used in performing indicator kriging as described in section 2-6. The kriging estimates essentially reflect a

30 Jun 97

probability that the concentration at the points of estimation exceed the clean-up value or other standard. These kriging estimates can be contoured to define areas or volumes of material that have a certain likelihood of exceeding some cleanup value. The contour value is essentially the probability of exceedance. Lastly, the size of the area defined by different probabilities of exceedance can be determined and, using a unit cost or similar approach, a cost-versus-risk curve can be developed. This can be used in programming money for the project, as a basis for negotiating cleanup levels with regulators, or to help determine if the cost and time of additional characterization work will be offset by less risk during construction. Alternatively, rather than transforming the data to ones and zeros, the actual values are kriged and the kriging variances can be used to determine prediction intervals on each estimated value as described in section 2-6. In the vicinity of the point estimate, these prediction intervals can be used to define the spread of potential values expected within a given probability. This assumes the data are normally distributed or have been transformed to be normally distributed.

6. Actions Required.

a. USACE elements identified in paragraph 2 shall consider applications of geostatistics as


FOR THE COMMANDER:

2 Appendices
App A - References
App B - Notation

described in this document as appropriate. This is particularly true during planning of large-scale site characterization efforts or when there are risk management or design decisions to be made that must consider the uncertainty of site characterization results. The same USACE elements should also encourage the use of geostatistics, where appropriate, by their contractors.

b. USACE elements shall make every effort to familiarize staff members actively supporting HTRW projects with the fundamentals and potential benefits of the application of geostatistics. This letter is a good starting point for learning about the use of geostatistics for HTRW projects. Users are encouraged to attend appropriate training.

c. This letter sets out procedures for the technically correct application of geostatistics which are consistent with current practice, such as set forth in ASTM D-5922 and D-5549. The technical procedures outlined herein shall be considered when performing USACE in-house geostatistical analysis or reviewing such analyses done by USACE contractors.


KISUK CHEUNG, P.E.
Acting Chief, Environmental Division
Directorate of Military Programs

CEMP-RT

**DEPARTMENT OF THE ARMY
U.S. Army Corps of Engineers
Washington, DC 20314-1000**

ETL 1110-1-175

Technical Letter
No. 1110-1-175

30 June 1997

**Engineering and Design
PRACTICAL ASPECTS OF APPLYING GEOSTATISTICS
AT HAZARDOUS, TOXIC, AND RADIOACTIVE WASTE SITES**

Table of Contents

Subject	Paragraph	Page	Subject	Paragraph	Page
Chapter 1			Chapter 4		
Introduction			Practical Aspects of Variogram		
General	1-1	1-1	Construction and Interpretation		
Scope	1-2	1-1	General	4-1	4-1
Organization	1-3	1-2	General Computation of		
An Overview of the Use of			Empirical Variogram	4-2	4-2
Geostatistics in Hazardous,			Nonstationarity	4-3	4-3
Toxic, and Radioactive Waste			Variogram Refinement	4-4	4-6
Site Investigations	1-4	1-2	Transformations and		
An Overview of Some Technical			Anisotropy Considerations . . .	4-5	4-9
Aspects of Geostatistics	1-5	1-4	Fitting a Theoretical		
			Variogram to the Sample		
Chapter 2			Variogram Points	4-6	4-10
Technical Aspects			Additional Trend		
of Geostatistics			Considerations	4-7	4-13
General	2-1	2-1	Outlier Detection	4-8	4-14
Regionalized Random Variables . .	2-2	2-1	Cross-Validation for		
Variograms	2-3	2-6	Model Verification	4-9	4-14
Kriging	2-4	2-9			
Co-kriging	2-5	2-14	Chapter 5		
Using Kriging to Assess Risk	2-6	2-15	Practical Aspects of		
			Geostatistics in Hazardous,		
Chapter 3			Toxic, and Radioactive		
Geostatistical Resources			Waste Site Investigations		
and Tools			General	5-1	5-1
Texts on Geostatistics	3-1	3-1	Water-Level Examples	5-2	5-1
Useful Journals	3-2	3-1	Bedrock-Elevation Examples . .	5-3	5-6
Software	3-3	3-2	Water-Quality Examples	5-4	5-13

Subject	Paragraph	Page	Subject	Paragraph	Page
Chapter 6					
Review of Kriging Applications			Simple Moving Average	7-3	7-2
Applicability of Kriging	6-1	6-1	Inverse-Distance Squared		
Important Elements of			Weighted Average	7-4	7-2
Kriging Applications	6-2	6-2	Triangulation	7-5	7-3
Errors in Measured Data	6-3	6-3	Splines	7-6	7-3
			Trend-Surface Analysis	7-7	7-5
			Simulation	7-8	7-5
Chapter 7					
Other Spatial Prediction			Appendix A: References		
Techniques			Appendix B: Notation		
General	7-1	7-1			
Global Measure of Central					
Tendency (Simple Averaging) . . .	7-2	7-1			

Chapter 1 Introduction

1-1. General

a. This Engineer Technical Letter (ETL) addresses the use of geostatistics at hazardous, toxic, and radioactive waste (HTRW) sites. One very fundamental aspect of perhaps all HTRW site investigations that deal with environmental contamination is the need to characterize the extent and spatial distribution of contamination. Such a characterization typically would include describing, using a variety of statistical or analytical tools, spatial trends and variability. A principal difficulty in doing this is the fact that measurements may be few, or may be sparsely scattered over large regions. A question that arises naturally in this situation is how one might interpolate in order to make predictions (or estimates) at points where measurements of contaminant concentration are not available. Such interpolation will be referred to as point, or punctual, estimation in this ETL. Additionally, an investigator may need to determine a single representative value for an area that is represented by several measured or estimated values or both; this will be referred to in this ETL as block estimation. Geostatistics is a set of statistical procedures designed to accomplish these ends. Geostatistics may be applied to many problems, other than contamination, that occur at HTRW sites. Even though this document addresses only twodimensional applications, geostatistics can be used in three dimensions as well. Indeed, there are many cases in which the third dimension, usually stratification, is desirable to address.

b. Kriging is the principal geostatistical methodology described in this ETL. For introductory purposes kriging can be defined as a technique for determining the optimal weighting of measurements at sampled locations for obtaining predictions, or estimates, at unsampled locations; additional definition of kriging is provided throughout this document. Kriging is well-suited for making point and block estimates. However, much of the advantage of using geostatistical procedures,

such as kriging, lies not just in the point and block estimates they provide, but in the information they provide concerning uncertainty associated with these estimates. The uncertainty information is usually quantified as either the standard deviation (or variance) associated with kriging estimates and is referred to as kriging standard deviation (or kriging variance) in this ETL.

c. Original geostatistical work involved making estimates for the areal extent and concentrations of economic mineral deposits, in relation to mining. Today (1996), geostatistical techniques continue to have a function in mining. However, a well-developed methodology that is capable of interpolating a given set of measured values at discrete locations into estimates for new locations or developing an individual estimate for an area including many locations, or both, has attracted users from many disciplines, and there is a trend toward incorporating geostatistics as standard curriculum for most geo-science educational programs. The use of geostatistical techniques as part of HTRW site investigations is becoming common because of the almost routine need for data interpolation as part of these investigations.

d. Once investigators have established that their data are adequate as to quality and quantity, geostatistics can provide powerful analytical tools that result in quantitative characterization of areas of special interest within the study area or the entire study area. These characterizations may address spatial variation; for example, it may be determined where values for concentrations of contaminants in soils are relatively high or low, are less than or greater than a specified value, or even have a high or low probability of exceeding a certain value.

1-2. Scope

a. The scope of this ETL will be limited principally to discussions and examples of two-dimensional point and block estimations using a geostatistical method known as kriging. The ETL will present the technical aspects of geostatistics

through discussion of the assumptions behind and the mechanics of several types of kriging, including ordinary kriging, which is applicable when the mean for the variable of interest is constant over the region of interest, and universal kriging, which is applicable when the mean for the variable of interest changes gradually over the region. The discussion also will address a specialized form of kriging known as indicator kriging and the use of information concerning uncertainty associated with kriging estimates. The fundamental concepts of geostatistical kriging theory will be provided in this ETL; however, references will be provided for additional and more detailed information.

b. The practical aspects of kriging will be discussed through categorical examples of HTRW site investigations. The phrase "HTRW site investigations," will refer to planning, analysis, and remediation implementation phases of HTRW projects.

c. Additional topics included in this ETL such as review of applications and of some of the newer geostatistical techniques will be limited. The intent will be to familiarize the reader with these topics and not to provide how-to knowledge.

1-3. Organization

a. This ETL is organized into seven chapters. Chapter 1 is introductory and includes an overview of the technical aspects of spatial prediction in general and certain geostatistical concepts. Chapter 2 provides a detailed discussion of assumptions and theory behind kriging, including equations and concepts that will be useful to investigators who wish to gain a better understanding of the technical aspects, or mathematics, of kriging interpolation. As indicated, many of the concepts developed in Chapter 2 are discussed in very general terms in Chapter 1, so those readers desiring only an overview of kriging concepts may wish to read only Chapter 1 and bypass Chapter 2 altogether.

b. Chapter 3 provides a review of texts that contain much more detailed information regarding

kriging theory than material included in Chapter 2. Chapter 3 also provides a brief generic discussion of kriging software.

c. Chapter 4 provides a detailed step-by-step discussion of variogram construction and demonstrates some pitfalls and solutions to this crucial process. Chapter 4 also discusses methodologies which investigators may use to evaluate their variograms.

d. Chapter 5 provides a discussion of practical aspects of geostatistics in a presentation of several example kriging applications with data from the HTRW field. The examples are intended to illustrate a few of the many different ways kriging can be used in HTRW site investigations and are not presented with the same level of detail used in Chapter 4.

e. Chapter 6 provides additional detail on some crucial aspects of kriging applications and includes considerations investigators may use to help determine if kriging is feasible for the application they have in mind, or reviewers can determine if the application of geostatistics was appropriate.

f. Chapter 7 provides an introduction to other methods for spatial modeling. This section also includes discussion of advanced stochastic methods such as simulation.

1-4. An Overview of the Use of Geostatistics in Hazardous, Toxic, and Radioactive Waste Site Investigations

a. General.

(1) HTRW site investigations involve complex administrative, scientific, and engineering functions and are truly interdisciplinary. Scientists and engineers, for instance, may be confronted with administrative findings or directives, associated with fiscal, managerial, or regulatory input, that may either guide or constrain their work. In a

likewise fashion, scientific findings may define the scope of administrative effort.

(2) Scientists and engineers involved in HTRW site investigations have found that they have an implicit need for many disciplines to fulfill the objectives of each particular investigation. Frequently, an HTRW site investigation will benefit from input from earth-science disciplines such as geology, hydrogeology, and chemistry, among others. Some HTRW site investigations are large enough to use several individuals from each of these disciplines, as well as many others, for the duration of multi-year investigations. Most disciplines associated with HTRW site investigations will benefit from knowledge or input from specialized and/or interdisciplinary branches; the geologist, for example, will occasionally benefit from knowledge of geophysics. Naturally, interdisciplinary input also can be very helpful, especially in geostatistics, where earth-science disciplines rely on assistance from statisticians.

(3) In this ETL and for its purposes, a complete HTRW site investigation is described concerning three relatively broad sequential activities or phases. These phases are referred to as initial planning, analysis, and implementation of remediation plans. Another very important HTRW site investigation activity, monitoring, is less discrete and is a part of all three phases. Monitoring represents the basis for analysis, is often modified as a result of analysis, and may be newly implemented as part of remediation.

(4) Kriging techniques can and have been used in any of the three phases. Only a few very basic applications of kriging techniques are described in this ETL. The intent of this ETL is to describe basic concepts so that more elaborate applications can be done based on a fundamental understanding of the procedures involved.

(5) For examples of more elaborate applications, the reader can refer to the material cited in Chapter 3. However, the best applications are developed by readers who have a clear understanding of the goals associated with each

particular HTRW site investigation and also have a good basic understanding of the fundamental geostatistical techniques. As alluded to here and elsewhere in this ETL, there are many techniques available for gridding data; kriging has an added advantage of generating kriging standard deviations that can be used as a measure of uncertainty.

b. Initial planning.

(1) Initial planning may involve several aspects associated with implementing or operating a monitoring network; it also may involve reconnaissance evaluation of an existing network. Additionally, because monitoring is present in all phases of HTRW site investigations, the same opportunities for geostatistical applications associated with network analysis that occur in the initial planning stages may occur, perhaps often, throughout the investigation. The information available from kriging standard deviations can add much to sampling or monitoring network analysis.

(2) For application of geostatistical techniques, the most likely aspects of network implementation and operation to be addressed certainly include network design, evaluation, and modification. Geostatistics offer the investigator opportunities to:

(a) Locate areas where existing sampling or monitoring networks may provide strong or weak estimates.

(b) Quantify the effect of increasing or decreasing the sampling or monitoring network density.

(c) Evaluate the effect of removing or relocating certain monitoring locations or adding new locations to the sampling or monitoring network.

c. Analysis.

(1) Although aspects of network design can be quite important during analysis, the investigator is likely to be concerned principally with using information from monitoring networks to evaluate

environmental conditions throughout the specified study area. The evaluations may require either point or block estimates. Often, design factors are addressed in the analytical phase as well.

(2) A common application for kriging techniques in HTRW site investigations is estimating real means. More common, however, is estimating the extent of areal contamination. Usually these estimates involve chemicals in air, water, and soil; however, if sufficient information is available, such estimates could include a wide range of environmental factors that involve many issues other than contaminants. Perhaps the most common examples concern geologic and hydrologic factors, such as depth to bedrock and groundwater-level elevations. Investigators need to realize that almost any set of measurements can be distributed using kriging techniques, providing there is a sufficient amount and distribution of measured information.

(3) The investigator also needs to realize that the resultant kriging estimates can be gridded. This gridding affords investigators opportunities to perform mathematical or logical operations, or both, on the kriging estimates, provided that investigators are comfortable with kriging estimates. Saturated thickness could, for example, be calculated from kriging estimates for groundwater elevations and base of aquifer elevations.

(4) Often, after preparing estimates for areal properties, the investigator may appreciate the opportunity afforded by kriging techniques to evaluate the confidence associated with the estimates. Maps of kriging standard deviations can provide the investigator with information concerning the confidence associated with the kriging estimates. Although the areas of lowest confidence may be well-known intuitively, maps of the kriging standard deviation are an important step toward quantification. More often than not, even the most experienced investigator will benefit from careful study of maps of kriging standard deviations.

d. Implementation of remediation.

(1) One of the most common applications for kriging techniques in the final phases of HTRW site investigations is evaluating compliance. For instance, a question such as "Is the mean concentration of constituent x within compliance limits?" is ubiquitous to HTRW site investigations. Making determinations concerning compliance is very similar to estimating areal extent as part of the analysis. Investigators and managers have much to gain from the confidence information available from kriging techniques as to the reliability of estimates as well as in optimizing monitoring networks.

(2) Kriging can also be very useful if managers are interested in making decisions based on the probability of certain conditions existing. If a condition can be defined by the manager, then, providing there are adequate data, indicator kriging can provide an estimate for the probability of existence. A common example of this kind of application is making areal determinations for probabilities that concentrations for a constituent do or do not exceed, for example, an action level.

(3) There are many operational remediation issues that kriging techniques may address as well. Remedial activities at HTRW sites often need estimates for amounts in general. For instance, there could be a need for information regarding volumes of contaminants to be treated, volumes of soil to be excavated, volumes of soil to be stored, and so on. By combining estimates for different geologic, hydrologic, and chemical factors, estimates for these volumes can be obtained from kriging techniques in much the same way as saturated thicknesses can be calculated.

1-5. An Overview of Some Technical Aspects of Geostatistics

The purpose of this section is to provide an overview of some of the procedures and concepts to be treated in detail in this ETL. Some of the technical ideas and terminology will be introduced in very general terms, with the goal of orienting the

reader who may not be familiar with the area of geostatistics.

a. General considerations in spatial prediction.

(1) The principal technical issue considered in this ETL is spatial prediction or modeling values of a spatial process; in particular it is considered how best to make use of measurements of a variable (such as pollutant concentration) at sampled locations to make inferences (or predictions) about that variable at unsampled locations or about values of the variable for the region as a whole.

(2) A spatial process can be viewed as having a large-scale or regional component and a smaller scale or local component; both of these components need to be accounted for when modeling a spatial process. The large-scale component is referred to as the mean field and is most often modeled by a spatial trend which may or may not be constant over the region. The smaller scale component is a random fluctuation which is mathematically combined with the trend to make up the sample at a point. The random component is usually assumed to be zero on the average but can be either positive or negative in individual samples. The separation of the trend from the random components is problem- and scale-dependent and requires some judgment to determine. There can be several "solutions" to the problem of separating the trend and random components that may be useful for various geostatistical purposes when using a single set of data.

(3) Local-scale fluctuation of the variable of interest (e.g. water levels or contaminant concentrations) at a sample point, although random, can show some association (i.e. correlation) with the random fluctuations at nearby points. This is referred to as spatial correlation. Positive spatial correlation between measurements means that the random components at both points tend to have the same sign, whereas negative correlation means the random components tend to have opposite signs. Both the "large-scale" trend and the positive spatial correlation of the "local-scale" fluctuations

contribute to measurements taken at locations close together being more closely related than measurements taken farther apart.

(4) The most obvious way one might proceed for spatial prediction at unsampled locations is simply to take an average of the sample values that one does have and assume that this value gives a reasonable prediction at all locations in the region of interest. This may work adequately in some cases, but one can also see the pitfalls in doing this. Using a single value for an entire region makes an implicit assumption of spatial homogeneity. It ignores any spatial trends that might exist in the data and it also ignores spatial continuity. If it is known that the variable of interest does have the tendency to be spatially correlated, then it would make sense to use a weighted average rather than a simple average in making a spatial prediction, with measurements at sampled locations that are nearer to the unsampled location being given more weight. This then is the motivation for the geostatistical methods discussed in this ETL. The method known as kriging, which is the principal subject to be considered here, is a technique for determining in an optimal manner the weighting of measurements at sampled locations for obtaining predictions at unsampled locations. These optimal weights depend on spatial trends and correlations that may be present.

(5) There are a number of ways to go about performing spatial prediction. The geostatistical method of kriging covered in this ETL belongs to a class of methods known as stochastic methods. In these methods, it is assumed that the measurements, both actual and potential, constitute a single realization of a random (or stochastic) process. One advantage of assuming the existence of such a random process is that measures of uncertainty, such as the variance used in kriging, can be defined. These measures of uncertainty permit objective assessment of the performance of a spatial prediction technique on the basis of how small such measures are. Once a measure of uncertainty has been selected, the weights to be used in spatial prediction may be determined so as to explicitly minimize the measure of uncertainty.

In short, the use of stochastic techniques provides the investigator with a way of objectively quantifying errors and determining weights. In practice, spatial predictions obtained using kriging are almost always accompanied by a measure of the associated error. Most kriging practitioners consider such an error evaluation to be an integral part of the analysis, and point to error analysis as one of the principal advantages of using kriging (or stochastic techniques in general) over other procedures.

(6) Nonstochastic techniques, on the other hand, are typically applied strictly empirically, with no assumptions concerning the existence of an underlying random process and with no theoretical framework with which to evaluate statistically the performance or optimality of the techniques. When they are applied in such a manner, it is not possible to evaluate in advance whether such a procedure would be expected to yield results that are satisfactory. Two techniques that are commonly applied in a nonstochastic setting are simple averaging, mentioned above, and trend analysis, which is a least-squares method for fitting a smooth surface to the data. Even though these techniques are usually applied nonstochastically, it is still possible to assess their performance if a stochastic setting is assumed. Loosely speaking (these ideas are discussed more precisely in Chapter 7), simple averaging would perform well if there is no trend and no spatial correlation, and trend analysis would perform well if there is a trend that can be modeled, but no spatial correlation. Lack of correlation in the observations is one assumption that is made in ordinary statistical regression analysis, and in fact trend analysis, if it is placed in a stochastic setting, is actually one special type of regression. The stochastic method of kriging explicitly incorporates the spatial correlations which are ignored in trend analysis. In Chapter 7, a few other common techniques that are usually applied in a nonstochastic setting will be discussed briefly. Most of these techniques are designed to incorporate the notion of spatial continuity, but the way it is incorporated may be subjective. Kriging provides an objective means of incorporating the presence of spatial correlation

and makes explicit the background assumptions that are being made.

b. Important geostatistical concepts. Below are some of the key ideas in geostatistics that will be given detailed attention in this ETL. They are introduced in much the same order that they are discussed in Chapter 2, where more detail is presented.

(1) Variograms.

(a) A central idea in geostatistics is the use of spatial correlation to improve spatial predictions, or interpolations. The variogram is the principal tool used to characterize the degree of spatial correlation present in the data and is fundamental to kriging. The correlation between measurements at two points is usually assumed, as described above, to depend on the separation between the two points. Values for all possible pairings of sample points can be examined by squaring the difference between the values in each pair. The squared differences are then categorized according to the distance separating the pair. For small separations, or lags, the squared differences are usually small and increase as the lag increases. A plot of the squared differences per sample pair as a function of lag is referred to as the sample variogram.

(b) The general behavior of the sample variogram points relates to the spatial correlation between sample sites and can provide investigators with qualitative information about the spatial process, but in order to use this information in a mathematically rigorous manner as a basis for interpolation, a function with specific properties must be fit to the sample variogram points. The fit, as with all curve-fitting procedures, takes the scattered points and passes a smooth curve through the points. The curve, which can be represented by a mathematical expression or function, is called a model. Several named models with characteristic features introduced in Chapter 2 are commonly used in geostatistics. The resultant variogram model is used to determine kriging weights for use in interpolation.

(2) Directional variogram and anisotropy. It is often the case that spatial correlation depends not only on distance between points, but also on direction. For example, measurements at pairs of points 100 m apart with the line between them oriented in a north-south direction may have a different correlation than measurements at points the same distance apart but with the line joining them oriented in an east-west direction. The spatial process is said to exhibit anisotropy, and what is known as a directional variogram must be used for the geostatistical analysis.

(3) Kriging and kriging variance.

(a) Kriging yields optimal spatial estimates at points where no measurements exist in terms of the values at points where one does have data. As discussed above, placing the problem in a stochastic framework permits precision-defining optimality. In kriging, the restriction is first imposed that the predicted value at any point is a linear combination of the measured values; that is, the kriging estimate is a linear predictor. Given this restriction, the values of the coefficients in this linear function are chosen so as to force the predictor to be optimal.

(b) The first criterion imposed is that the estimate be unbiased, or that in an average sense the difference between the predicted value and actual value is zero. The second optimality criterion is that the prediction variance be minimized. This variance is a statistical error measure defined to be the average squared difference between predicted and actual values. Because the kriging estimate minimizes this variance, it is known as the best (minimum variance) unbiased linear predictor. This minimization is performed algebraically and results in a set of equations known as the kriging equations, which give an explicit representation of the optimal coefficients (weights) in terms of the variogram. The form of these equations is presented in Chapter 2.

(c) Also given in Chapter 2 is an expression for the kriging variance. This variance depends on geometry of the data sites, with the variance at

locations near points with measurements tending to be smaller. One can then associate with any spatial prediction a variance, which gives an indication of the uncertainty in that predicted value. As mentioned before, this measure of uncertainty gives kriging one of its principal advantages over many other techniques.

(4) Trends and universal kriging. Special attention must be given in kriging to the question of whether there are spatial trends in the data. A trend in this case is usually any detectable tendency for the measurements to change as a function of the coordinate variables but can also be a function of other explanatory variables. For example, aside from random fluctuations, measurements of groundwater elevations may exhibit a tendency to increase in a consistent manner the farther one proceeds in a certain direction. A kriging analysis in which there is no spatial trend is known as ordinary kriging; when a trend does exist, universal kriging should be considered. In universal kriging, one attempts to account for the trends present. For example, it might be assumed that the trend can be represented as a linear function of coordinate variables. The form of the trend model is then incorporated into the universal kriging equations to obtain the optimal weights.

(5) Block kriging. What has been discussed in the preceding paragraphs is usually known as point, or punctual, kriging. In point kriging, the goal is to predict the value of a variable at discrete locations. By contrast, in block kriging the goal is to predict the average value, over a specified region, of a variable. As in point kriging, the optimal predictor is a linear combination of the measured data values, and degree of uncertainty is indicated by a block kriging variance. Block kriging variances tend to be smaller than point kriging variances because averages tend to be less variable than individual values.

(6) Prediction intervals and normality.

(a) A standard kriging analysis will give two values for any location: the optimal kriging estimate and the kriging variance. The variance

provides a measure of uncertainty for the prediction. In some cases, it may be desirable to go even further in specifying the nature of the uncertainty than simply giving the variance. One way to proceed is to try to obtain what is known as a prediction interval. Here one seeks an interval such that there is a certain probability, typically 95 percent, that the actual value lies in this interval.

(b) Finding such an interval often hinges on having knowledge of the probability distribution of the variables being sampled. One ideal situation is when the variable of interest, e.g., contaminant concentration, can be assumed to have a normal distribution. In this case, given the set of measured values, a potential value at an unsampled location has a normal distribution with mean given by the kriging estimate and variance given by the kriging variance. It is thus, using classical statistics, straightforward to use this normal distribution to obtain a 95 percent prediction interval for concentration at the unsampled location.

(7) Transformations. Having a prediction interval will generally be much more informative than simply having the kriging estimate and kriging variance, which explains why investigators often ask whether normality assumptions can be made for their data. When a normality assumption cannot be made, it is sometimes possible to find a transformation that will make the data normal, or nearly so. For example, a transformation that is often tried is the logarithmic transformation. That is, one simply takes the logarithm of all data values (assuming they are > 0) and performs the geostatistical analysis on these transformed values rather than on the original data. Prediction intervals obtained using transformed values can be readily converted to corresponding intervals on untransformed variables. There are, however, subtleties that must be considered in back-transforming the kriging estimate and the kriging variance; these are discussed in more detail in Chapter 2.

(8) Indicator kriging.

(a) In indicator kriging, analysis is performed using what are known as indicator variables rather than the measured data themselves. An indicator variable is thus a special kind of transform of the measured data and can have only two possible values: 0 or 1. To obtain the indicator variables to be analyzed, first specify a threshold value, say c , which may represent, for example, a contaminant concentration level which is of particular importance. At each measurement location, the indicator variable is then assigned a value of 1 if the measured value is less than or equal to c , and is assigned a value of 0 if the measured value is greater than c . This kind of transform will allow censored data, or data reported as less than some reporting limit, to be included in the analysis if the reporting limit is less than or equal to the cutoff value of c . After the indicator transform has been performed, the kriging analysis is performed using these indicator variables in the same manner discussed above; first a variogram is obtained, and the kriging equations yield the optimal linear predictor and the kriging variance for the indicators.

(b) Whereas the indicator kriging analysis is done using only 0's and 1's, the interpolated estimates are not restricted to these two values. In most cases the estimates are between 0 and 1, which is interpreted to be the probability that the actual value is less than or equal to the threshold c . Performing this analysis for a number of different threshold values, c , can give the investigator information about the probability distribution of contaminant values at a location, which may in turn be used to obtain prediction intervals. As discussed above, such intervals may even be more valuable than having only the optimal predictor and variance provided by the usual kriging analysis, particularly if behavior of extremes may be of interest to the investigator. The advantage of using indicator kriging to obtain prediction intervals is that it is not necessary to assume a distribution for the data, as in the discussion of normality above.

Chapter 2 Technical Aspects of Geostatistics

2-1. General

a. This chapter provides technical aspects or the necessary theoretical background for understanding kriging applications. Emphasis will be placed on presentation of the basic ideas; long formulas or derivations are kept to a minimum. Statistical terms that are commonly used in geostatistical applications will be highlighted with bold text and briefly defined as they are introduced; notation used in this ETL is also tabulated in Appendix B. The reader who wishes a more thorough discussion of these fundamental concepts may consult the references cited in Chapter 3. Previous exposure to engineering statistics at the level of Devore (1987) and Ross (1987) would be helpful in understanding some parts of this chapter. Readers with limited statistical experience may wish to briefly scan this chapter and refer back to it after reading the remaining chapters.

b. In section 2-2, regionalized random variables are discussed. Regionalized random variables constitute the random process that is sampled to obtain the observed data that are available for analysis. Basic ideas related to probability distributions, means, variances, and correlation are introduced. The variogram, which is the fundamental tool used in geostatistics to analyze spatial correlation, is introduced in section 2-3. In section 2-4 how kriging is used to obtain the best weights for spatial prediction is discussed, and how the mean squared prediction error for these predictions is computed is also shown. Section 2-5 deals briefly with co-kriging, which is prediction of one variable based not only on measurements of that variable but on measurements of other variables as well. Finally, section 2-6 shows how kriging may be applied to determine not just optimal spatial predictions but also probabilities associated with various events, such as extreme events that may be of importance in risk-based analyses.

2-2. Regionalized Random Variables

a. General.

(1) Suppose the extent of groundwater contamination of a particular pollutant over a given study area is being determined. To simplify the presentation, all data are assumed to be distributed over a two-dimensional region. In three-dimensional groundwater flow systems, one could study the depth-averaged concentration of a pollutant or the concentration of the pollutant in a particular horizontal stratum of the flow system. Let a vector $\underline{x}=(u,v)$ denote an arbitrary spatial location in the study area. Unless otherwise stated, it will be assumed throughout the ETL that u is the east-west coordinate and v is the north-south coordinate (Figure 2-1). Denote by $z(\underline{x})$ a measurement at location \underline{x} , such as the concentration of a pollutant. The ultimate goal of an investigator would be to determine $z(\underline{x})$ for all locations in the study area. However, without explicit knowledge of the flow and transport field, this goal cannot be achieved. Therefore, suppose, instead, that the goal is to estimate the values of $z(\underline{x})$ with a given error tolerance. In other situations, small estimation error over some parts of the study area (for instance, near a domestic water supply) may need to be obtained, while allowing larger estimation errors in other parts of the study area. The theory of regionalized random variables is designed to accomplish these goals.

(2) In the regionalized random variable theory, the true measurement $z(\underline{x})$ is assumed to be the value of a **random variable** $Z(\underline{x})$. Associating a random variable $Z(\underline{x})$ with a true measurement $z(\underline{x})$ is done for the purpose of characterizing the degree of uncertainty in the quantity of interest at point \underline{x} . If there is no actual measurement taken at \underline{x} , then the values taken on by $Z(\underline{x})$ represent "potential" measurements at \underline{x} ; that is, $Z(\underline{x})$ represents possible values that might be expected if a measurement were taken at \underline{x} . Because there is uncertainty associated with $Z(\underline{x})$, it needs to be characterized by a **probability distribution**, defined by $P[Z(\underline{x}) \leq c]$ where P denotes probability and c is any constant.

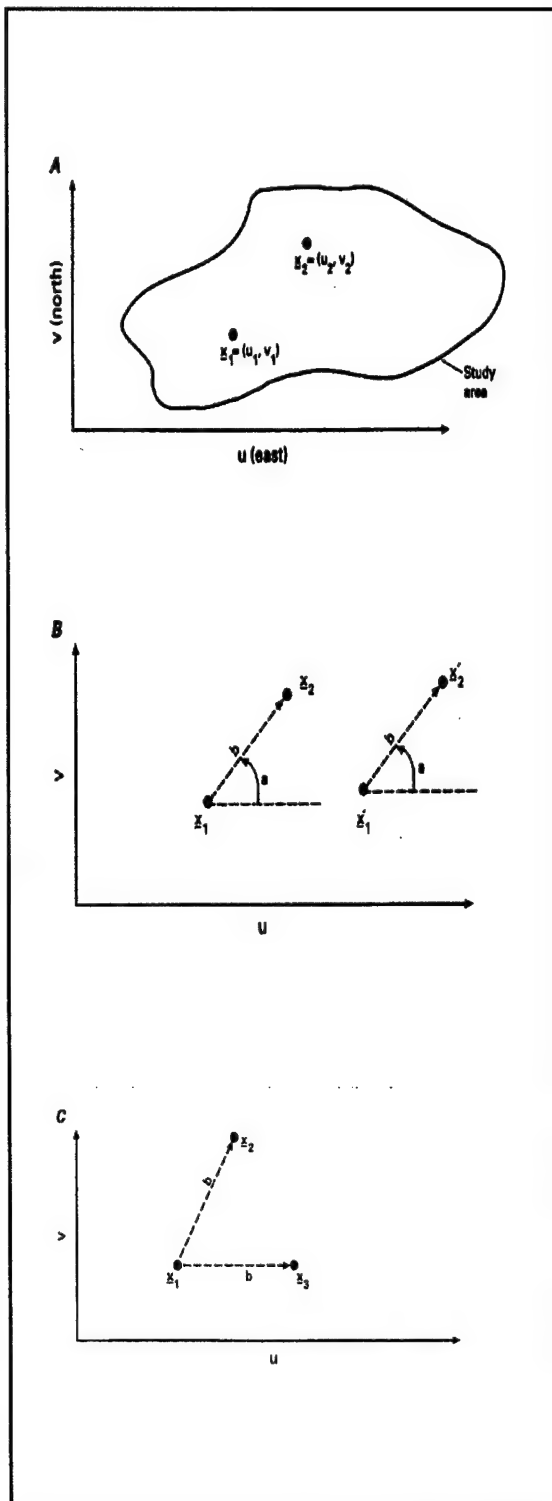


Figure 2-1. Diagrams showing A, hypothetical study area; B, stationary covariance function; and C, isotropic covariance function

This distribution is a function of c , and, to be completely defined, needs to be known for all values of c . The distribution is used to make evaluations such as: suppose that we have no measurement of concentration of a certain contaminant at \underline{x} , but the distribution is known, and a threshold value of $c = 8$ mg/l is of interest. If $P[Z(\underline{x}) \leq 8] = 0.60$, then, if a measurement were made at \underline{x} , there is a 60-percent chance of obtaining a value less than or equal to 8 mg/l. The distribution also may be used to calculate other probabilities, such as the probability of obtaining a value in some specified interval.

(3) An important concept to keep in mind in all geostatistical applications is the **support** of the regionalized random variable. The support of $Z(\underline{x})$ is the in situ geometric unit represented by an individual sample. For example, in a soil contamination study, sample $Z(\underline{x})$ might represent the concentration of a contaminant in a vertical soil core 0.1 m in diameter and 1 m in length, and centered at location \underline{x} . Thus, even though $Z(\underline{x})$ is defined at a particular point, it is representative of a volume of soil. Changing the support of $Z(\underline{x})$ will usually change its probability distribution. Therefore, the observations in a geostatistical analysis should all have the same support. The method called point, or punctual, kriging, described in section 2-4, is designed to predict values of $Z(\underline{x})$ with the same support as the sample data.

(4) A concept closely related to support is that of **estimation block**, which is a geometric unit larger than the support of a single observation, for which a single representative value is desired. For example, in the above soil contamination study, it may be necessary to estimate the average concentration of the contaminant in a truckload of soil excavated from a block 6 m long, 6 m wide, and 0.3 m thick. Using a method called block kriging, also described in section 2-4, the block average can be predicted based on individual measurements.

(5) Although the distribution of $Z(\underline{x})$ completely characterizes $Z(\underline{x})$ at any particular location, this distribution indicates nothing about the relations among the values of $Z(\underline{x})$ at different

locations, which is very important, because geostatistics is based on using a measurement of a regionalized variable at one location to gain information about values of the variable at another location. The notion of distribution of $Z(\underline{x})$ at a single location is readily generalized to two or more locations. For two locations, if we let \underline{x}_1 and \underline{x}_2 be two distinct locations, then the **joint probability distribution** is defined to be the probability $P [Z(\underline{x}_1) \leq c_1, Z(\underline{x}_2) \leq c_2]$ for any constants c_1 and c_2 . This latter probability means the probability that both $Z(\underline{x}_1) \leq c_1$ and $Z(\underline{x}_2) \leq c_2$. If the variables $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$ are statistically independent of one another, then the joint probability distribution can be obtained as the product of the individual probability distributions,

$$\begin{aligned} P [Z(\underline{x}_1) \leq c_1, Z(\underline{x}_2) \leq c_2] \\ = P [Z(\underline{x}_1) \leq c_1] P [Z(\underline{x}_2) \leq c_2] \end{aligned} \quad (2-1)$$

However, in most applications, $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$ will not be statistically independent and their joint distribution cannot be obtained from the individual distributions. When this joint distribution description is applied to more than two locations, specification of the full spatial distribution of Z would require knowing the joint distribution of $Z(\underline{x}_1), \dots, Z(\underline{x}_n)$ for any set of n spatial locations and for any n ; however, except in very special cases, working with the full set of distribution functions of $Z(\underline{x})$ is not feasible and is not done.

(6) To simplify the problem even further, various parameters of the distributions are usually considered rather than dealing with the entire distributions. The parameter most commonly used to characterize a distribution is the **mean**, or, because the mean in geostatistical applications depends on the spatial variable \underline{x} , the mean may be called the **spatial mean**, or the drift. In statistics, the mean is referred to as the expectation (E) of the random variable $Z(\underline{x})$, and the symbol m is used in this report to denote this expectation. Thus,

$$\mu(\underline{x}) = E [Z(\underline{x})] \quad (2-2)$$

is used to denote the mean, or expected value, of the bracketed term, in this case $Z(\underline{x})$. It is intuitively helpful to think of the expectation as an average. In fact, if the distribution of $Z(\underline{x})$ assigned equal probability to a finite number of values, then the expectation of $Z(\underline{x})$ would indeed be the simple average of these numbers. In geostatistics, however, $Z(\underline{x})$ is usually assumed to take on any value in a continuous range of possible values, rather than being limited to a discrete set of values. In this case, calculus needs to be used to define the expectation. The following example illustrates the difference between averages and expectations.

b. Example 1.

(1) An experiment consists of injecting a conservative tracer at a particular well in a steady-state groundwater flow system and measuring the concentration, $Z_1(\underline{x})$, of the tracer in a neighboring well 24 hr later. The tracer is then allowed to flush from the system, and the experiment is repeated a second time to obtain another concentration measurement, $Z_2(\underline{x})$, at the same location. If this process is repeated n times, n concentration measurements $Z_1(\underline{x}), Z_2(\underline{x}), \dots, Z_n(\underline{x})$ would be obtained, all at location \underline{x} . The average concentration level at location \underline{x} is

$$\begin{aligned} \bar{Z}_n(\underline{x}) = \frac{1}{n} (Z_1(\underline{x}) + Z_2(\underline{x}) \\ + \dots + Z_n(\underline{x})) \end{aligned} \quad (2-3)$$

which would change depending on n and on the actual values obtained for $Z_1(\underline{x}), Z_2(\underline{x}), \dots, Z_n(\underline{x})$. However, in the limit as n increases, $\bar{Z}_n(\underline{x})$ becomes closer and closer to the true mean, or expected, concentration $\mu(\underline{x})$:

$$\bar{Z}_n(\underline{x}) \rightarrow \mu(\underline{x}) \text{ as } n \text{ increases} \quad (2-4)$$

This theoretical limit is a constant value, or **population parameter**, as opposed to $\bar{Z}_n(\underline{x})$, which is a random variable, or a property of the particular sample that is taken.

(2) In example 1, no assumptions were needed concerning whether the mean changed with spatial location, because all sampling was done at one sampling location \underline{x} . In most HTRW applications, the mean will probably change depending on the sampling location. In addition, usually only one observation is available at any particular location. Therefore some assumptions regarding the structure of $\mu(\underline{x})$ must be made. For example, it is sometimes appropriate to assume $\mu(\underline{x}) = \mu$ is constant for all \underline{x} , in which case $Z(\underline{x})$ is said to have a **stationary mean**. Data which have no underlying trend such as hydraulic conductivity in a homogeneous aquifer, for example, might be assumed to have a constant mean. If the mean is constant, it makes sense to estimate it with the sample average of n observations taken at different spatial locations $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$

$$\bar{Z}_n = \frac{1}{n} [Z(\underline{x}_1) + Z(\underline{x}_2) + \dots + Z(\underline{x}_n)] \quad (2-5)$$

However, in contrast to example 1, \bar{Z}_n defined in this way may not get closer to μ as n gets large. Because of the possible spatial correlation in the data, the size of the sampling region must be large in relation to the correlation length in order for \bar{Z}_n to accurately estimate μ .

(3) In addition to the mean of $Z(\underline{x})$, its variability or dispersion is also of interest, and this variability is most commonly measured by the **(spatial) variance**, defined to be the mean of squared deviations of $Z(\underline{x})$ from $\mu(\underline{x})$ and denoted by $\sigma^2(\underline{x})$.

$$\sigma^2(\underline{x}) = E[(Z(\underline{x}) - \mu(\underline{x}))^2] \quad (2-6)$$

The **(spatial) standard deviation** $\sigma(\underline{x})$ is the square root of the variance. The following example illustrates the difference between the population variance, which has been defined above, and a sample variance.

c. Example 2.

(1) If the scenario presented in example 1 is again used, the sample variance $S_n^2(\underline{x})$ of the n measurements could be computed as follows:

$$S_n^2(\underline{x}) = \frac{1}{n-1} \sum_{i=1}^n (Z_i(\underline{x}) - \bar{Z}_n(\underline{x}))^2 \quad (2-7)$$

This number gives a measure of dispersion of the $Z_i(\underline{x})$ values from their sample mean. The sample variance depends on n and on the particular values observed for $Z_1(\underline{x}), Z_2(\underline{x}), \dots, Z_n(\underline{x})$. However, in the limit as n increases, $S_n^2(\underline{x})$ gets closer and closer to a constant value, which is denoted by $\sigma^2(\underline{x})$. In this case, $\sigma^2(\underline{x})$ is a population parameter, and $S_n^2(\underline{x})$ is a random variable.

(2) The mean and variance can both be calculated from the probability distribution of $Z(\underline{x})$. Again, in geostatistics, the relations among regionalized variables at different locations are of interest. From the joint distribution of $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$ the **(spatial) covariance function**,

$$C(\underline{x}_1, \underline{x}_2) = E[(Z(\underline{x}_1) - \mu(\underline{x}_1))(Z(\underline{x}_2) - \mu(\underline{x}_2))] \quad (2-8)$$

may be obtained. This function has a key role in geostatistical analyses. It is a measure of association between values obtained at point \underline{x}_1 and those obtained at point \underline{x}_2 . If values at these two spatial locations tend to be greater than average or less than average at the same time, then the covariance will be positive. However, if the values vary in the opposite direction (that is, one tends to be larger than average when the other is less than average, and vice versa), the covariance will be negative.

(3) Because $C(\underline{x}_1, \underline{x}_2)$ is an unknown population parameter, it too must be estimated using a statistic computed from sample data. To make this possible, it is often assumed that the covariance function depends only on the distance between points, which is defined as the **lag** h , and not on their relative location or orientation,

$$C(\underline{x}_1, \underline{x}_2) = C(h), \quad (2-9)$$

$$h = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}$$

Under this assumption, $C(h)$ can be estimated by pooling all pairs of observations that are approximately h units apart and computing a **sample covariance function**

$$\begin{aligned} \hat{C}(h) = & \text{average} \left\{ \left(Z(\underline{x}_i) - \bar{Z} \right) \left(Z(\underline{x}_j) - \bar{Z} \right) : \right. \\ & \left. h - \Delta h < h_{ij} < h + \Delta h \right\} \end{aligned} \quad (2-10)$$

where h_{ij} is the distance between \underline{x}_i and \underline{x}_j and the average is over all pairs of points such that h_{ij} is between $h - \Delta h$ and $h + \Delta h$. The distance h is called the lag and Δh is called the **lag tolerance**. There are more effective ways to estimate $C(h)$ other than using Equation 2-10; for example, see Isaaks and Srivastava (1989). However, because the emphasis in this ETL is on the variogram (to be defined below) rather than the covariance function, we will not need to use the estimated covariance function.

(4) A covariance function is called **stationary** if it does not depend on the origin of the coordinate system, that is,

$$C(\underline{x}_1 + \underline{b}, \underline{x}_2 + \underline{b}) = C(\underline{x}_1, \underline{x}_2) \quad (2-11)$$

for any given vector, \underline{b} (Figure 2-1). The covariance function (Equation 2-9) is stationary because changing the origin does not change the distance between the points. Substituting $\underline{x}_1 = \underline{x}_2 = \underline{x}$ in Equation 2-9 yields

$$C(\underline{x}, \underline{x}) = C(0) \quad (2-12)$$

which, combined with the definitions in Equations 2-6 and 2-8, becomes

$$\sigma^2(\underline{x}) = C(0) \text{ for all } \underline{x} \quad (2-13)$$

Therefore, when $Z(\underline{x})$ has a stationary covariance function, the variance of $Z(\underline{x})$ is constant for all \underline{x} . The covariance function can then be standardized by dividing it by the variance. The resulting dimensionless function of h is called the **spatial correlation function**,

$$\rho(h) = \frac{C(h)}{C(0)} \quad (2-14)$$

The correlation function is a scale-independent measure of linear association between values of Z at different locations. The spatial correlation is always between -1 and +1, with a value of zero indicating no linear association.

(5) In addition to being stationary, the covariance function in Equation 2-9 has another important property. It is also **isotropic**, or **omni-directional**, because it does not depend on the direction between the two locations. In many HTRW applications, the correlation between values of Z at two locations is a function of direction as well as lag. For example, contaminant concentrations in a groundwater flow system might be more highly correlated along a transect in the direction of flow than along a transect perpendicular to the flow. In that case, the covariance function depends on both the lag h and the angle a between locations,

$$C(\underline{x}_1, \underline{x}_2) = C(h, a),$$

$$h = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}, \quad (2-15)$$

$$a = \text{atan} \left(\frac{v_2 - v_1}{u_2 - u_1} \right)$$

Here, a is the angle measured counterclockwise from the east direction (Figure 2-1). In many geostatistical publications or computer packages, the angle may be defined as clockwise from the north direction, so care should be taken in defining the appropriate angle in any application. A covariance function satisfying Equation 2-15 is called **anisotropic**, or **multi-directional**.

(6) To summarize, the basic model framework that will be used throughout the ETL is the following: the value of a measurement $z(\underline{x})$ (concentration, porosity, hydraulic head, and so on) at location \underline{x} of a two-dimensional region is the value of a regionalized random variable, $Z(\underline{x})$, with mean $\mu(\underline{x})$ and stationary covariance function $C(h, a)$. Other assumptions may be added in the applications sections to analyze specific data sets, but this framework will be the basic framework from which many of the results will be derived. In some situations, the covariance stationarity assumption may be relaxed, for instance, when using the linear variogram described in the next section.

2-3. Variograms

a. Regionalized random variables differ from classical (ordinary least-squares) regression models in that the **residuals**, defined as the deviations of the regionalized random variable from its mean and denoted by

$$Z^*(\underline{x}) = Z(\underline{x}) - \mu(\underline{x}) \quad (2-16)$$

are related to one another, whereas the residuals in a regression model are generally assumed to be independent. Thus, in the regionalized random-variable model, observed values of the residuals from sampled locations contain valuable information when predicting the value of $Z(\underline{x})$ at unsampled sites. The relationship among the residuals can be understood by examining the variogram, which is a tool that is widely used in geostatistics for modeling the degree of spatial dependence in a regionalized random variable. Although the variogram is closely related to the covariance function, there are some important differences between the variogram and covariance function that will be described below. The covariance function, and related correlation function, are more commonly used in basic statistics courses than the variogram, so many readers may be more familiar with the former concepts. However, the variogram is more widely used in geostatistics, and because of this we

will adopt the variogram as the primary tool for analyzing spatial dependence in the remainder of this ETL.

b. As was the case with the covariance function, it is necessary to distinguish between the theoretical variogram, which is a population parameter, and the sample variogram, which is an estimator of the theoretical variogram obtained from observed data. The **theoretical variogram** of a regionalized random variable, $\gamma(\underline{x}_1, \underline{x}_2)$ is defined as one half of the variance of the difference between residuals at locations \underline{x}_1 and \underline{x}_2 :

$$\gamma(\underline{x}_1, \underline{x}_2) = \frac{1}{2} \text{Var} [Z^*(\underline{x}_1) - Z^*(\underline{x}_2)] \quad (2-17)$$

Because the residuals have been mean-centered, as shown in Equation 2-16, they have a mean of zero. Therefore, using the well-known formula for the variance of a random variable X

$$\text{Var}(X) = E(X^2) - (EX)^2 \quad (2-18)$$

it is seen that Equation 2-17 is equivalent to

$$\gamma(\underline{x}_1, \underline{x}_2) = \frac{1}{2} E [Z^*(\underline{x}_1) - Z^*(\underline{x}_2)]^2 \quad (2-19)$$

The theoretical variogram is always non-negative, with a small value of g indicating that the residuals at locations \underline{x}_1 and \underline{x}_2 tend to be close and a large value of λ indicating that the residuals tend to be different. Equation 2-19 is sometimes called a **semi-variogram**, because of the multiplication by $1/2$, but will be referred to in this ETL as a variogram.

c. It would be ideal to know the theoretical variogram before taking observations, but unfortunately, it must be estimated using sample data. To facilitate variogram estimation, it is usually assumed in a similar manner to the covariance function that γ depends only on the lag,

$$\gamma(x_1, x_2) = \gamma(h), \quad (2-20)$$

$$h = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}$$

or possibly, on the lag and angle between locations

$$\gamma(x_1, x_2) = \gamma(h, a),$$

$$h = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}, \quad (2-21)$$

$$a = \text{atan} \left(\frac{v_2 - v_1}{u_2 - u_1} \right)$$

(Figure 2-1). Equation 2-20 is called an **isotropic variogram** and Equation 2-21 is a **directional variogram** at angle a .

d. For the isotropic case, the **sample**, or **empirical, variogram** is obtained by averaging the square of all computed differences between residuals separated by a given lag:

$$\hat{\gamma}(h) = \frac{1}{2} \text{ave} \left\{ \left(Z^*(x_1) - Z^*(x_2) \right)^2 : \right. \quad (2-22)$$

$$h - \Delta h < h_{ij} < h + \Delta h$$

where, as before, h_{ij} is the distance between x_i and x_j . For a given h as more and more points separated by distance $h \pm \Delta h$ are sampled and as Δh gets small, $\hat{\gamma}(h)$ should approach the theoretical variogram. More detail on variogram estimation will be presented in Chapter 4, including the directional case. In this section, it will be sufficient to describe some general properties of isotropic variograms that will be referred to numerous times in the application sections to follow.

e. A plot of the sample variogram versus h often has a considerable degree of scatter (Figure 2-2), which is especially evident if the sample size n is small. However, the points can usually be

fitted by a smooth curve that represents a theoretical variogram selected from a suite of possible choices. Usually, the theoretical variogram is monotonically increasing, signifying that the farther two observations are apart, the more their residuals tend to differ, on average, from one another. Several properties common to many theoretical variograms are shown in Figure 2-2. If the variogram either reaches or becomes asymptotic to a constant value as h increases, that value is called the **sill** (Figure 2-2). The distance (value of h) after which the variogram remains at or close to the sill is called the **range**. Measurements whose locations are farther apart than the range all have the same degree of association. Often, a variogram will have a discontinuity at the origin, signifying that even measurements very close together are not identical. Such variation in the measurements at small scales is called the **nugget effect**. The size of the discontinuity is called the nugget. Although the nugget effect is sometimes confused with measurement error, there is a subtle difference between these two concepts that will be explained in section 2-4. A simple monotonic function is usually selected to approximate the variogram. Four such functions that are often used in practice are:

the **exponential variogram** (parameters: sill, $s > 0$; nugget, $0 < g < s$; range, $r > 0$)

$$\gamma(h) = \begin{cases} g + (s - g) \left[1 - \exp \left(-3 \frac{h}{r} \right) \right], & h > 0 \\ 0, & h = 0 \end{cases} \quad (2-23)$$

the **spherical variogram** (parameters: sill, $s > 0$; nugget, $0 < g < s$; range, $r > 0$)

$$\gamma(h) = \begin{cases} s, & h > r \\ g + (s - g) \left[1.5 \frac{h}{r} - 0.5 \left(\frac{h}{r} \right)^3 \right], & 0 < h \leq r \\ 0, & h = 0 \end{cases} \quad (2-24)$$

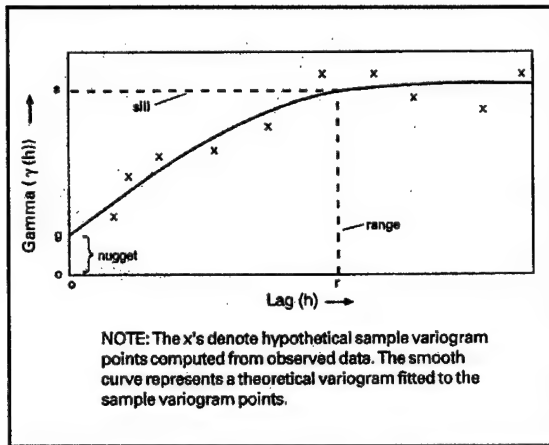


Figure 2-2. Diagram showing variogram and features

the **Gaussian variogram** (parameters: sill, $s > 0$; nugget, $0 < g < s$; range, $r > 0$)

$$\gamma(h) = \begin{cases} g + (s-g) \left[1 - \exp \left(-3 \left(\frac{h}{r} \right)^2 \right) \right], & h > 0 \\ 0, & h = 0 \end{cases} \quad (2-25)$$

and, the **linear variogram** (parameters: nugget, $g > 0$; slope, $b > 0$)

$$\gamma(h) = \begin{cases} g + bh, & h > 0 \\ 0, & h = 0 \end{cases} \quad (2-26)$$

f. Although there are many other models that are used for variograms (Journel and Huijbregts 1978), these four are the most commonly used and are shown in Figure 2-3. The exponential, spherical, and Gaussian models are similar in that they all have a sill and a range. However, they have different shapes near zero lag ($h=0$) that, as will be discussed in Chapter 4, result in significant differences in the prediction results using the three models. The linear model is quite different from the other three, in that it does not reach a sill, but increases linearly without. This fact will have important implications on the prediction results using a linear variogram. Because the squared differences between residuals tend to increase

without bound as the lag increases, a regionalized random variable with a linear variogram will have ever-increasing variability about its mean as the size of the sampling region is increased. In applications involving the linear variogram, the variogram is usually truncated at a sill corresponding to the value of the variogram at maximum lag h_{\max} .

g. Before closing this section, it will be useful to highlight some similarities and contrasts between the covariance function and the variogram. Although the variogram is commonly used in a geostatistical analysis, it is sometimes easier to gain an intuitive understanding of the methodology using the covariance function, or equivalently, the spatial variance and the correlation function. When $Z(x)$ has a stationary, isotropic covariance function (Equation 2-9), there is a one-to-one correspondence between the variogram and the covariance function, namely

$$\gamma(h) = C(0) - C(h) \quad (2-27)$$

As long as $C(h)$ approaches zero as h increases (a minor technicality that can always be assumed in practice), then, as indicated by Equation 2-27, the variogram reaches a sill and the sill equals $C(0)$. Therefore, when dealing with a covariance-stationary regionalized random variable, the variogram and the spatial covariance function contain the same information as one another. By factoring out $C(0)=s$ from Equation 2-27 and using Equation 2-14, the relationship between the spatial correlation function and the variogram can be obtained

$$\rho(h) = 1 - \frac{\gamma(h)}{s} \quad (2-28)$$

From Equation 2-28, it is evident that high values of $\gamma(h)$ (i.e., close to s) signify low values of $\rho(h)$. In fact, $\rho(h) = 0$ whenever $\gamma(h) = s$, indicating that observations whose locations are farther apart than the range are uncorrelated. As h gets small, a nugget in $\gamma(h)$ is reflected in a correlation that is less than 1

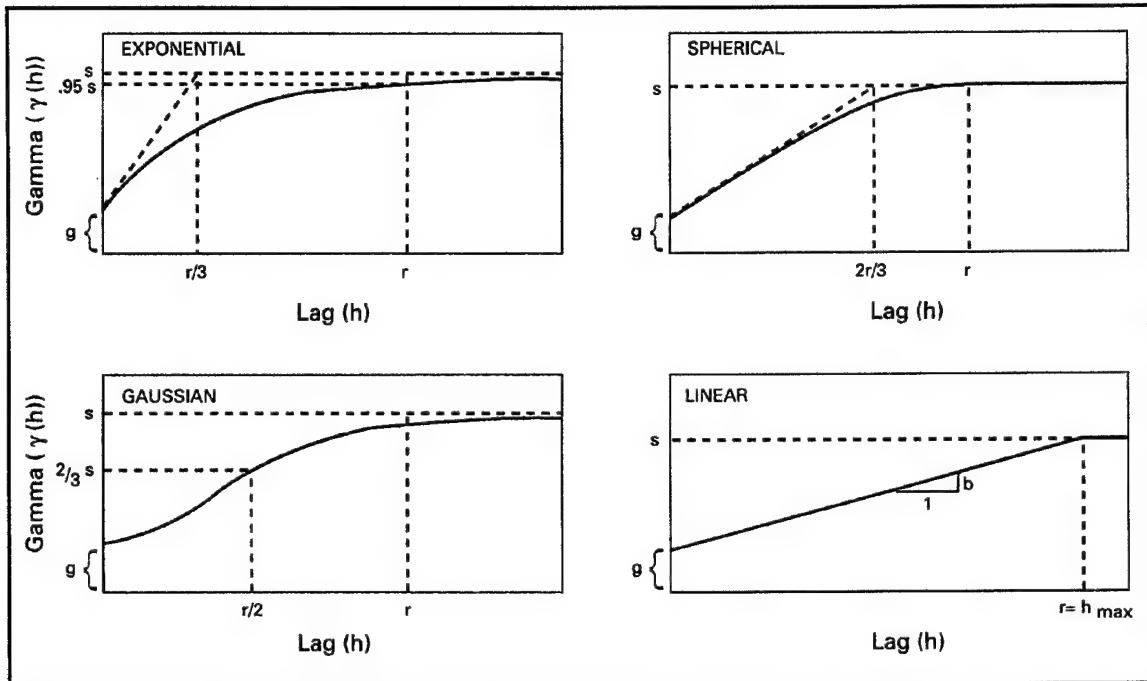


Figure 2-3. Theoretical variograms showing A, exponential; B, spherical; C, Gaussian; and D, linear models

$$\rho(h) \rightarrow 1 - \frac{g}{s} \text{ as } h \rightarrow 0 \quad (2-29)$$

Therefore, the larger g is in relation to s , the less correlated nearby observations are. The case when $g=s$, called a **pure nugget variogram**, results in $\rho(h)=0$ for all $h>0$. In that case, neighboring observations are uncorrelated no matter how closely they are spaced.

h . Occasionally, $\gamma(h)$ may not reach a finite sill, as in the linear variogram Equation 2-26. In that case, it is not possible to define a correlation function as in Equation 2-28. The corresponding regionalized random variable is said to be **intrinsically stationary** (Journel and Huijbregts 1978), which is more general than covariance stationarity. The theory behind intrinsically stationary variograms will not be presented in this ETL. As long as a "pseudo-range" h_{\max} is defined, all of the computations described below can be generalized.

2-4. Kriging

a. General.

(1) Given a regionalized random variable $Z(x)$ with a known theoretical variogram, the question is: how can the value of $Z(x)$ be predicted at an arbitrary location, based on measurements taken at other locations? Suppose that Z is measured at n specified locations: $Z(x_1), \dots, Z(x_n)$. For example, Z could correspond to hydraulic conductivity and the locations might correspond to n preexisting wells in an aquifer. Let a new location be given by $x_0=(u_0, v_0)$ and denote the i th measurement location by $x_i=(u_i, v_i)$. Suppose that, based on prior knowledge of the geology, there are no prevailing trends in hydraulic conductivity, so the mean of $Z(x)$ is assumed to be constant over the entire region:

$$\mu(x) = \mu \text{ (constant)} \quad (2-30)$$

(2) Suppose the investigator wants to predict the value of $Z(x_0)$ by using a **linear predictor**, $\hat{Z}(x_0)$, which is defined as a weighted linear combination of the measured data

$$\hat{Z}(x_0) = \sum_{i=1}^n w_i Z(x_i) \quad (2-31)$$

where w_i is the weight assigned to $Z(x_i)$. To determine specific values for the weights, some criteria need to be specified for $\hat{Z}(x_0)$ to be a good predictor of $Z(x_0)$. The first criterion is that $\hat{Z}(x_0)$ be an **unbiased predictor** of $Z(x_0)$, which is expressed as

$$E[\hat{Z}(x_0) - Z(x_0)] = 0 \quad (2-32)$$

(3) An unbiased predictor will neither consistently overpredict nor underpredict $Z(x_0)$ because the statistical expectation of the prediction errors is zero. The second criterion for a good predictor is that it have small **prediction variance**, defined by

$$\begin{aligned} \text{Var}[\hat{Z}(x_0) - Z(x_0)] \\ = E[(\hat{Z}(x_0) - Z(x_0))^2] \end{aligned} \quad (2-33)$$

(4) The smaller the prediction variance, the closer $\hat{Z}(x_0)$ will be (on average) to the true value $Z(x_0)$. The geostatistical method of kriging deals with computing the **best linear unbiased predictor** of $Z(x_0)$, which is the linear unbiased predictor (Equations 2-31 and 2-32) with the smallest possible prediction variance (Equation 2-33).

(5) The form of the best linear unbiased predictor will depend on the mean of $Z(x)$. For example, if $Z(x)$ has a constant mean (Equation 2-30) and a pure nugget variogram [$\gamma(h)=s$ for all $h>0$], the best linear unbiased predictor of $Z(x_0)$ will simply be the average of the measured data

$$\hat{Z}(x_0) = \frac{1}{n} \sum_{i=1}^n Z(x_i) \quad (2-34)$$

Because the variogram is the same for all $h>0$ and there is no trend in the data, there is no reason to favor any of the measurements over any of the other measurements. Therefore, the weights are all the same. Ordinary kriging, which is discussed in section 2-4b, deals with the constant-mean model (assumption in Equation 2-30) in which the variogram is not a pure nugget variogram. The weights of the best linear unbiased predictor will reflect the information in the variogram and will result in an improved predictor over the sample mean. In section 2-4c, universal kriging, which is the extension of ordinary kriging to the case of a nonconstant mean, is discussed. Universal kriging is a very powerful tool that can be used to combine regression models and spatial prediction into one unifying theory. Other, more specialized types of kriging that will be discussed in this section are indicator kriging (section 2-6c), block kriging (section 2-4d), and co-kriging (section 2-5).

(6) Before giving the kriging equations, one final note is in order. There is a prediction technique in geostatistics known as **simple kriging**, which deals with best linear unbiased prediction in the case when the mean of $Z(x)$ is fixed and known. Simple kriging is not discussed in this ETL, because, in most applications, the mean is not known and has to be estimated.

b. Ordinary kriging.

(1) General.

(a) Let $Z(x)$ be a regionalized random variable with constant mean (Equation 2-30) and isotropic variogram (Equation 2-20). Also, assume that the variogram reaches a sill so that the variance of $Z(x)$ is $C(0)=s$, and the correlation function is given by Equation 2-28. Although the prediction equations can be expressed in terms of the variogram, they will be defined here in terms of the sill (variance) and the correlation function.

(b) Consider linear unbiased predictors of the form of Equation 2-31 with the condition in Equation 2-32 holding. The unbiased condition is

equivalent to $\mu \sum_{i=1}^n w_i = \mu$ for any μ , which holds

if and only if $\sum_{i=1}^n w_i = 1$. Therefore, all linear

unbiased estimators need to have weights that sum to one. There are many sets of weights that satisfy this condition, including the set with all the weights equal to $1/n$, as in the sample mean, Equation 2-34. However, the unique set of weights that minimize the prediction variance (Equation 2-33) can be shown to satisfy the following set of $n+1$ **ordinary kriging equations** (Chapter 12, Isaaks and Srivastava (1989)):

$$\sum_{j=1}^n w_j \rho_{ij} + \frac{\lambda}{s} = \rho_{i0}, \quad i=1, 2, \dots, n, \quad (2-35a)$$

$$\sum_{j=1}^n w_j = 1 \quad (2-35b)$$

where $\rho_{ij} = \rho(h_{ij})$ is the correlation between observations i and j , h_{ij} is the distance between locations i and j , and λ is a coefficient resulting from the constrained optimization. Furthermore, the resulting **ordinary kriging variance** is

$$\begin{aligned} \sigma_k^2(x_0) &= E \left[\left(\hat{Z}(x_0) - Z(x_0) \right)^2 \right] \\ &= s \left(1 - \sum_{j=1}^n w_j \rho_{j0} \right) - \lambda \end{aligned} \quad (2-36)$$

(c) The system of Equations 2-35a and 2-35b can easily be solved for the w_i 's and λ , after which the kriging variance can be obtained from Equation 2-36. Note that the ordinary kriging variance changes depending on the prediction location x_0 , even though the variance of $Z(x_0)$ itself (Equation 2-6) is constant for all x_0 .

(2) Example 1.

(a) Let the mean of $Z(x)$ satisfy Equation 2-30, and suppose that the residual $Z^*(x)$ (Equation 2-16) has an isotropic exponential variogram (Equation 2-23). Consider predicting $Z(x_0)$ based

on $n=2$ measurements $Z(x_1)$ and $Z(x_2)$, where the three locations (x_0 , x_1 , and x_2) are distinct. Using Equations 2-23 and 2-28, note that the correlation function is

$$\rho(h) = \begin{cases} \left(1 - \frac{g}{s} \right) \exp \left(-3 \frac{h}{r} \right), & h > 0 \\ 1, & h = 0 \end{cases} \quad (2-37)$$

For illustrative purposes, suppose that

$$\frac{g}{s} = p, \quad 0 \leq p \leq 1 \quad (2-38)$$

where p is a fixed proportion. The quantity p is sometimes referred to as a **relative nugget**.

(b) The ordinary kriging Equations 2-35a and 2-35b are given by

$$w_1 + w_2 \rho_{12} + \frac{\lambda}{s} = \rho_{10} \quad (2-39a)$$

$$w_1 \rho_{12} + w_2 + \frac{\lambda}{s} = \rho_{20} \quad (2-39b)$$

$$w_1 + w_2 = 1 \quad (2-39c)$$

These three equations have three unknowns: w_1 , w_2 , and λ ; the solution is

$$w_1 = \frac{1}{2} + \frac{1}{2} \frac{\rho_{10} - \rho_{20}}{1 - \rho_{12}} \quad (2-40a)$$

$$w_2 = \frac{1}{2} - \frac{1}{2} \frac{\rho_{10} - \rho_{20}}{1 - \rho_{12}} \quad (2-40b)$$

and

$$\lambda = \frac{s}{2} (\rho_{10} + \rho_{20} - \rho_{12} - 1) \quad (2-41)$$

The resulting kriging variance is

$$\sigma_K^2(\underline{x})_0 \quad (2-42)$$

$$= s \left[\frac{3}{2} - w_1 \rho_{10} - w_2 \rho_{20} - \frac{1}{2} (\rho_{10} + \rho_{20} - \rho_{12}) \right]$$

Although there are only three sample locations in this example (two actual and one potential), it indicates several properties of best linear unbiased prediction that hold in general. For example,

(c) **Effect of sill.** The kriging weights depend on s only through the relative nugget p . However, the kriging variance is directly proportional to s . The sill is called a scaling parameter because scaling each measurement by a constant c has the effect of scaling s by c^2 . When the relative nugget is allowed to vary so that s and g can change independently, the effect of s is somewhat more complicated.

(d) **Effect of nugget.** Increasing p has the effect of drawing each of the weights closer to $1/2$. In fact, as p approaches 1, both weights will equal $1/2$. The larger g is in relation to s , the more small-scale variability there is in the data and the less important the correlation between neighboring locations becomes. The increased small-scale variability also causes an increase in the kriging variance.

(e) **Effect of correlations.** If $Z(\underline{x}_0)$ is more highly correlated with $Z(\underline{x}_1)$ than with $Z(\underline{x}_2)$, then w_1 will be larger than w_2 , indicating that the measurement at the first location has more predictive information than the measurement at the second location. Also, correlation in the data always decreases the kriging variance compared to the variance with uncorrelated data.

(f) **Effect of data clumping.** If $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$ are highly correlated, as indicated by ρ_{12} being close to 1, then the two measurements contain much of the same information. Two situations can occur: $\rho_{10} = \rho_{20}$, in which case the weights are both equal, or $\rho_{10} > \rho_{20}$ [$\rho_{10} < \rho_{20}$], in which case w_1 will be much larger [smaller] than w_2 . In either

case, the kriging variance will increase to reflect the redundant information in the two measurements. Automatic adjustment of the kriging weights and kriging variance to account for data clumping is an important property of the kriging predictor.

(3) Example 2 (Nugget effect versus measurement error).

(a) In example 1, all three locations \underline{x}_0 , \underline{x}_1 , and \underline{x}_2 , were assumed to be distinct. When a prediction location happens to coincide with a measurement location, there is an important distinction that needs to be made between a true nugget effect and a measurement error. Suppose that in example 1, \underline{x}_0 and \underline{x}_1 are the same. If there is only small-scale variability, but no measurement error, then repeated measurements at the same location should be identical, that is, $\rho_{10} = 1$. In this case, the kriging equations result in $w_1 = 1$, $w_2 = 0$, and $\lambda = 0$ and in a kriging variance of zero. That is, $Z(\underline{x}_1)$ is a perfect predictor of $Z(\underline{x}_0)$. This property, called **exact interpolation**, is a property of kriging when the data are assumed to contain no measurement errors. However, suppose instead that the nugget is interpreted as measurement error rather than small-scale variability. In that case, repeated measurements at the same location would not be perfectly correlated, but rather, $\rho_{10} = 1 - g/s$.

(b) Substituting this correlation into the kriging equations and solving the equations results in a predictor that does not exactly interpolate the data, but instead smooths the measured data to account for the measurement error. In this ETL, prediction locations are assumed not to coincide with measurement locations, in which case no distinction needs to be made between nugget and measurement error.

c. Universal kriging.

(1) Universal kriging is an extension of ordinary kriging, that, due to the fact that environmental data often contain drift, can be important in HTRW site investigations. Universal kriging addresses the case of a nonconstant mean $\mu(\underline{x})$.

Generally, the mean is assumed to have a functional dependence on spatial location of the form

$$\mu(u, v) = \sum_{j=1}^p \beta_j f_j(u, v) \quad (2-43)$$

where the $f_j(u, v)$'s are known deterministic functions of $\underline{x}=(u, v)$ (that is, these functions serve as independent variables) and the β_j 's are regression coefficients to be estimated from the data. For example, suppose $Z(\underline{x})$ is hydraulic head in an aquifer. If the flow is in a steady state, it might be reasonable to assume, in a given case, that the mean of $Z(\underline{x})$ has a unidirectional groundwater gradient that is given by

$$\mu(u, v) = \beta_1 + \beta_2 u \quad (2-44)$$

In this example, there are two independent variables:

$$\begin{aligned} f_1(u, v) &= 1 \\ f_2(u, v) &= u \end{aligned} \quad (2-45)$$

and two regression coefficients (β_1 and β_2). The mean can include other independent variables besides simple algebraic functions of u and v . For example, if the aquifer is not of uniform thickness, an independent variable that involves the aquifer thickness at location (u, v) could be included.

(2) The form assumed for the mean in Equation 2-43 is also generally used in standard linear regression analysis. In regression, ordinary least-squares is generally used to solve for the coefficients; when this is done, it is assumed that the residuals are independent and identically distributed. Universal kriging is an extension of ordinary least-squares regression that allows for spatially correlated residuals. Assuming that $Z(\underline{x})$ is a regionalized random variable with a mean as in Equation 2-43 and residual correlation function as in Equation 2-28, the best linear unbiased predictor (Equation 2-10) can be obtained from the following $n+p$ equations, called the **universal kriging equations** (Journel and Huijbregts 1978):

$$\begin{aligned} \sum_{j=1}^n w_j \rho_{ij} + \frac{1}{s} \sum_{k=1}^p \lambda_k f_k(\underline{x}_i) \\ = \rho_{i0}, \quad i=1, 2, \dots, n \end{aligned} \quad (2-46a)$$

$$\begin{aligned} \sum_{j=1}^n w_j f_k(\underline{x}_j) \\ = f_k(\underline{x}_0), \quad k=1, 2, \dots, p \end{aligned} \quad (2-46b)$$

where, in contrast to the ordinary kriging equations (2-35a and b), there are now p coefficients $\lambda_1, \dots, \lambda_p$ resulting from the unbiased condition on the predictor. The first term in the mean (Equation 2-43) will usually be a constant, or intercept, for which $f_1(\underline{x}) = 1$. Therefore, the universal kriging model includes ordinary kriging as a special case. The **universal kriging variance** is given by

$$\begin{aligned} \sigma_k^2(\underline{x}_0) &= s \left(1 - \sum_{i=1}^n w_i \rho_{i0} \right) \\ &\quad - \sum_{k=1}^p \lambda_k f_k(\underline{x}_0) \end{aligned} \quad (2-47)$$

These equations can be easily solved to obtain universal kriging predictors and kriging variances for any desired location. The estimated trend surface does not actually need to be computed to obtain the universal kriging predictor. If a particular application needs an estimate of the trend surface, then generalized least-squares regression can be used to estimate the coefficients (β_j 's) in the regression equation.

d. Block kriging.

(1) Up to this point, the problem of predicting the value of a regionalized random variable at a given location in the region over which the variable is defined has been considered. Implicit in this analysis is the assumption that the support of the variable being predicted is defined in exactly the same way as the variables that make up the measurements. However, there may be applications

where it is necessary to estimate the average value of Z over an estimation block of much larger area than is represented by an individual sample. For example, an estimate of the average concentration of a contaminant over an entire aquifer based on point measurements at various locations might be needed. In other applications, an estimate of the average concentration of soil contaminant in daily excavation volumes that are much larger than the volume of an individual sample may be needed. Let Z_B be the average value of $Z(x)$ over a particular block B ,

$$Z_B = \frac{1}{m} \sum_{i=1}^m Z(x_{0i}) \quad (2-48)$$

where x_{0i} , $i=1, \dots, m$, denotes m prediction locations in block B . The object is to predict this average rather than the regionalized variable at a single location. In many applications, the locations x_{0i} might correspond to nodes of a regular grid or finite-element nodes in a groundwater model. Results of the block kriging are dependent on m and on the placement of the prediction locations. Selecting a large number of locations in block B , where each location has approximately the same representative area, is the best approach (Chapter 13, Isaaks and Srivastava (1989)).

(2) The objective of block kriging is to obtain the best linear unbiased predictor of Z_B and an estimate of the block kriging variance based on the measurements. The model for $Z(x)$ can be the constant-mean model (Equation 2-30) assumed for ordinary kriging or the more general linear regression model (Equation 2-43) assumed for universal kriging. In either case, the predicted value of Z_B coincides with the average of the predicted values of the individual measurements in the block; that is

$$\hat{Z}_B = \frac{1}{m} \sum_{i=1}^m \hat{Z}(x_{0i}) \quad (2-49)$$

In this equation, the individual predicted values are obtained from either the ordinary or universal kriging equations. However, computation of the block

kriging variance is not as simple, because the individual kriging estimates are not independent of one another. There are simple modifications to the kriging equations discussed in sections 2-4b and 2-4c that can be used to directly compute the kriging estimate of Z_B , along with its kriging variance (Chapter 13, Isaaks and Srivastava (1989)). The equations are not presented in this ETL. The computer packages described in the next section can be used to compute block kriging estimates. In general, kriged values of block averages are less variable than kriged values at single locations. Consequently, the blocked kriging variance tends to be smaller than the kriging variance at a single location.

2-5. Co-kriging

a. Kriging as discussed so far provides a way of predicting values of a regionalized variable $Z(x)$ at a location x_0 based on measurements of the same variable at locations x_1, x_2, \dots, x_n . In some situations, however, there will be available measurements not only of $Z(x)$, but also of one or more other variables that can be used to improve predictions of $Z(x_0)$. The variable $Z(x)$ will be called the primary variable, because it is the one to be predicted, and the other variables will be called secondary variables. **Co-kriging** is the technique that allows the use of the information contained in secondary variables in the prediction of a primary variable. As an example, suppose that $Z(x)$ is a regionalized variable representing the hexavalent chromium concentration, a relatively difficult determination, and that hexavalent chromium concentration needs to be predicted at a location x_0 based on measurements of hexavalent chromium at other locations, but there are also measurements of a second relatively easily determined contaminant, for example lead, that tend to be correlated with hexavalent chromium concentration and these data are to be used as well. Denote the second variable lead by a regionalized variable $W(x)$, and assume that measurements have been made on W at m locations x'_1, x'_2, \dots, x'_m . The co-kriging predictor of $Z(x_0)$ is then

$$\hat{Z}_C(x_0) = \sum_{i=1}^n w_i Z(x_i) + \sum_{j=1}^m w'_j W(x'_j) \quad (2-50)$$

This is a straightforward extension of the kriging predictor in Equation 2-31. Analogous to kriging, co-kriging produces the weights w_i and w'_j so that the resulting predictor is the best linear unbiased predictor. Also, as with kriging, co-kriging requires modeling of the variogram for Z , but co-kriging presents the investigator with the additional necessity of modeling the variogram of W and the **cross variogram** for Z and W . The optimal weights are then expressed in terms of all these variogram properties. More than one secondary variable may be included in the co-kriging predictor, and theory has been developed for co-kriging in the presence of drift (universal co-kriging) and block co-kriging. Details are not included in this ETL, but the interested reader may refer to Isaaks and Srivastava (1989) and Deutsch and Journel (1992) for more discussion and citation of other references.

b. One situation in which co-kriging might be useful is when the primary variable is undersampled, so any additional information, such as that given by secondary variables, would be helpful. However, although co-kriging can be a useful tool, joint modeling of several variables tends to be demanding in terms of data and computational requirements. Thus, undersampling of the primary variable may present problems for co-kriging as well as for one-variable kriging. Also, unless the primary variable of interest is highly correlated with the secondary variable(s), the weights assigned to the secondary variable(s) are often small, with the result that the effort needed to include the additional variable(s) may not be worthwhile. For these reasons, co-kriging tends not to be used extensively in practice.

c. Although co-kriging is similar to universal kriging, in that both techniques use extra variables to help predict $Z(x)$, there is an important distinction between the two techniques. In

universal kriging, the independent variables in Equation 2-43 need to be known with certainty at the prediction location x_0 . For example, aquifer thickness might be an independent variable in predicting aquifer head if it can easily be determined at any location. However, aquifer thickness may need to be considered a secondary variable in a co-kriging procedure if it is only known at a few selected points in the aquifer.

2-6. Using Kriging to Assess Risk

a. General.

(1) The kriging predictor of $Z(x_0)$ has certain desirable properties with respect to how close it is to the actual value of $Z(x_0)$, it is unbiased and has smallest variance among all linear predictors. On the average, or in an expected sense, the predicted value will be near the actual value. When possible, however, the investigator would like to go further in specifying the relationship between the predicted and observed values. Ideally, the investigator would like to make probability statements. For example, if $Z(x_0)$ is concentration of a contaminant, the investigator might like to be 95 percent certain that the true concentration is within 0.05 ug/l of the predicted concentration. In other situations, the probability that the actual concentration exceeds a given target value might need to be estimated. Knowledge of the entire distribution function of $Z(x)$, as opposed to knowledge of only its mean and variogram, can be used for risk-qualified inferences in situations when extremes might be of more interest than averages.

(2) Introduction of the concept of a **conditional probability distribution** function of the regionalized variable $Z(x)$ is appropriate at this point. This concept will also be used in Chapter 7 when conditional simulation is discussed. The conditional probability distribution function has a definition much like that of the probability distribution function in section 2-2, except the probability that $Z(x) \leq c$ is computed "conditional on," or "given," information at other spatial locations. The interest in geostatistics is to make predictions

at a location x_0 using information at measurement locations x_1, x_2, \dots, x_n , so, in terms of conditional distributions, interest focuses on $P[Z(x_0) \leq c \mid Z(x_1), Z(x_2), \dots, Z(x_n)]$. The vertical bar denotes the conditioning and is read "given." This conditional probability distribution needs to be determined to make probability statements about the regionalized variable at location x_0 . Also, **conditional mean** and **conditional variance** can be defined in the present context in the same way that mean and variance for distribution functions were defined in section 2-2.

(3) Section 2-6b contains methods for using kriging output to obtain prediction intervals or quantiles when the regionalized random variable is either normally distributed or can be transformed to a near-normal distribution. Section 2-6c discusses indicator kriging, which is a nonparametric method for obtaining quantiles when data cannot be transformed adequately to a normal distribution.

b. Normal distributions and transformations.

(1) For prediction at a location x_0 , a kriging analysis produces the predictor $\hat{Z}(x_0)$ and the associated kriging variance $\sigma_k^2(x_0)$. If more informative probability assessments are to be made, the ideal situation is when $Z(x)$ can be assumed to be a Gaussian, or normal, process, which means that $[Z(x_1), \dots, Z(x_n)]$ has a joint normal probability distribution for any set of n locations and any value of n . In this case, the conditional probability distribution of $Z(x_0)$ given the n observations is a normal distribution with conditional mean equal to the kriging predictor $\hat{Z}(x_0)$ and conditional variance equal to the kriging variance $\sigma_k^2(x_0)$. This normal distribution can be used to obtain a **prediction interval** for $Z(x_0)$ (conditional on the measured data). For example, from a table of the normal distribution, a value of 1.96 corresponding to a 0.95 (two-sided) probability can be obtained. Then the assertion that there is a 95-percent chance that $Z(x_0)$ will be in the 95-percent prediction interval $[\hat{Z}(x_0) - 1.96 \sigma_k(x_0), \hat{Z}(x_0) + 1.96 \sigma_k(x_0)]$ can be made. Knowing this interval is much more useful than simply knowing the kriging predictor and variance.

(2) To illustrate quantile estimation, suppose that contaminant concentrations are being studied and the concentration that has only a 1-percent chance of being exceeded at location x_0 needs to be determined. The appropriate (one-sided) value from a normal table is 2.33, so the desired estimate is $\hat{Z}(x_0) + 2.33 \sigma_k(x_0)$.

(3) Even if $Z(x)$ is not Gaussian, it is often possible to find a **transformation**, $Y(x) = T(Z(x))$, such that $Y(x)$ is approximately Gaussian. When a transformation is made, the kriging analysis is performed using the transformed data $Y(x)$, and the inverse transformation may be applied to obtain prediction intervals for the original data. For example, the most common transformation is the (natural) **logarithmic transformation**, in which $Y(x) = \ln[Z(x)]$. A 95-percent prediction interval for $Z(x)$ is then $\{\exp[\hat{Y}(x_0) - 1.96 \sigma_k(x_0)], \exp[\hat{Y}(x_0) + 1.96 \sigma_k(x_0)]\}$. As long as the transformation is a one-to-one function such as a logarithmic transform, prediction intervals for the original data can be obtained by simply back-transforming prediction intervals for the transformed data.

(4) Although it is a simple matter to obtain prediction intervals and probabilities using simple back-transformation, it is more difficult to obtain a predictor of the untransformed data that is both unbiased and optimal in some sense. For example, in the case of a logarithmic transformation, a kriging analysis using the transformed data yields a predictor $\hat{Y}(x_0)$, which is the best linear unbiased predictor of $Y(x_0)$. However, the back-transformed value $\hat{Z}(x_0) = \exp[\hat{Y}(x_0)]$ does not possess these same optimality properties as a predictor of $Y(x_0)$. The methodology known as log-normal kriging, and more generally trans-normal kriging, has been developed to obtain predictors in this setting (Journel and Huijbregts 1978), but because of the complexity involved in these procedures, they are not usually used by practitioners. If a predicted value corresponding to $Z(x_0)$ needs to be obtained for purposes of contour plotting, the kriging predictions $\hat{Y}(x_0)$ may be back-transformed and plotted, as long as the investigator realizes that such values do not have the usual kriging optimality properties.

c. *Indicator kriging.*

(1) There may be situations when a transformation that makes $Z(x)$ approximately normal cannot be easily determined. In such situations, **indicator kriging** can be used to make inferences about the probability distribution of $Z(x)$. Because no distributional assumptions are made, this technique is known as a **nonparametric** statistical procedure. An example of indicator kriging is included in Chapter 5, and a paper by Journel (1988) is a good reference for additional information about indicator kriging.

(2) To perform indicator kriging, a special transformation, known as an indicator transformation, is applied to $Z(x)$:

$$I(x, c) = \begin{cases} 1, & Z(x) \leq c \\ 0, & Z(x) > c \end{cases} \quad (2-51)$$

If, as in the usual kriging scenario, the data set at hand consists of measurements of the regionalized variable $Z(x)$ at n locations, c needs to be fixed first, and then the indicator transformation is applied by replacing values that are less than or equal to c with 1 and values that are greater than c with 0. The variogram and kriging analysis is then performed using these 0's and 1's rather than the raw data.

(3) Kriging predictors using the indicator data will be equal to their observed values of 0 or 1 at the measurement locations x_i , $i=1, \dots, n$. However, at locations different from the measurement locations, predictions may be between 0 and 1. In interpreting these values, the power of indicator kriging becomes apparent. A predicted value at x_0

is an estimate of the conditional probability distribution $P[Z(x_0) \leq c | Z(x_1), Z(x_2), \dots, Z(x_n)]$. This analysis may be performed for a range of values of c , and by doing this the entire distribution function can be estimated. This estimate of the distribution function can be used in the same manner discussed above to obtain prediction intervals or estimates of quantiles. For example, to estimate the value that has a 1-percent chance of being exceeded at location x_0 , the value of c for which the kriged indicator prediction is 0.99 at that location is determined.

(4) One advantage of indicator kriging is that the indicator variogram is robust with respect to extreme outliers in the data because no matter how large (or small) $Z(x)$ is, the indicator variable is either 0 or 1. Indicator variables may also be used in the context of block kriging. For example, a spatial average of $I(x, c)$ over a block B equals the fraction of block B for which $Z(x)$ is less than c . Another advantage of indicator kriging is that it can be used when some data are censored.

(5) Despite the relative ease of implementation, there are several drawbacks to indicator kriging, and investigators may wish to use this technique only when other methods, such as normality transformations, produce unacceptable results. For example, the kriged values of $I(x, c)$ may be less than 0 or larger than 1. Also, the kriged prediction for $I(x, c_1)$ may be larger than the kriged prediction for $I(x, c_2)$ even if $c_1 < c_2$, which is not compatible with a valid probability distribution. There are several more advanced techniques for dealing with these problems (Chapter 18, Isaaks and Srivastava (1989); however, they are beyond the scope of this ETL.

Chapter 3 Geostatistical Resources and Tools

Since the mid-1970's, a myriad of texts and articles have been published that are either totally dedicated to geostatistical methods or discuss geostatistics in detail. Numerous computer programs and software packages on geostatistics and kriging accompany many of these texts. Although only a few of these resources will be briefly described in this ETL, their lists of references can provide the interested reader a path to other geostatistical topics or software not specifically covered in the resources.

3-1. Texts on Geostatistics

a. The geostatistical texts presented in this section can be classified into two broad categories: instructional texts or reference texts. For one who is delving into geostatistics for the first time, Clark's (1979) book is a starting point. Simple explanations of the basic kriging techniques are applied to an example data set. A more advanced treatment of the kriging techniques is described by Isaaks and Srivastava (1989). This textbook presents a detailed discussion of many of the background statistical tools and concepts needed in geostatistical applications, including histograms and distributions (univariate and bivariate), sampling, correlation, and spatial continuity. The text also discusses how to treat the subtleties of kriging using three data sets as examples. As well as being instructional, the book also can be used as a reference.

b. Texts by Cressie (1991) and Journel and Huijbregts (1978) describe the tools of geostatistics, but also include a comprehensive theoretical background on the techniques. Cressie's (1991) text is a treatment of spatial processes in general and reviews a wide range of statistical techniques in the analysis and stochastic modeling of spatial data. There is a four-chapter section on geostatistics, with a complete discussion of variogram estimation, kriging (including universal kriging),

intrinsic random functions, and comparisons of kriging to other spatial prediction techniques. The text is written from a statistician's point of view and is, in places, written at a fairly high level mathematically. It nevertheless contains numerous examples and illustrations using real-world data. Journel and Huijbregts (1978) maintain a mining-geological perspective. Two other texts written by statisticians that present general treatments of spatial processes, but that lack detailed discussions of kriging, are Cliff and Ord (1981) and Ripley (1981).

c. David's (1977) text was the first extensive discussion of geostatistics and kriging in mining applications, and the discussion is presented from a practitioner's viewpoint. Its value as reference material derives from the many specific mining applications and results. A broad statistics text with a bent toward geological applications (Davis (1986), serves as a reference for standard statistical procedures needed in geological applications of geostatistics. A book by Bras and Rodriguez-Itrube (1985) that discusses a range of techniques for stochastic modeling in the field of hydrology includes a chapter on applications of kriging. There is a fairly complete mathematical development of kriging with details of an application to predict mean areal precipitation. In a paper prepared for the U.S. Environmental Protection Agency, Journel (1993) discusses geostatistics as it relates to environmental science. Finally, Olea (1991) presents a useful glossary of geostatistical terms.

3-2. Useful Journals

The journal *Mathematical Geology* by the International Association for Mathematical Geologists reports new developments in the theory and application of kriging. Although many of the articles present new applications of kriging tools, many also are dedicated to the derivation of statistical properties of the variogram, kriging estimation, and cross-validation results. Journals such as *Water Resources Research*, published by the American Geophysical Union, and *Groundwater*,

published by the Association of Groundwater Scientists and Engineers, contain articles describing special applications of kriging techniques in the environmental arena. *Water Resources Research* tends to contain articles that are highly theoretical. Other journals that may contain information addressing geostatistics are the *Journal of Environmental Engineering*, published by the American Society of Civil Engineers; *Stochastic Hydrology and Hydraulics*, published by Springer International, and the *North American Council on Geostatistics*, published by the Colorado School of Mines.

3-3. Software

a. The geostatistics software described in this section is limited to a few readily available public domain packages that are executable at least on the DOS and sometimes on the UNIX platforms. There are several commercial packages that are being marketed, but these will not be reviewed in this ETL. It is beyond the scope of this ETL to acquire and evaluate commercial packages; however, a matrix-like table (Table 3-1) has been included. The table addresses each of the software packages described in this ETL and also may serve as a reference guide to other software packages.

b. Some of the earliest interactive kriging software offered as a package was developed by Grundy and Miesch (1987). Overall, this general statistics package (STATPAC) contains a series of programs that can handle two-dimensional kriging, including universal kriging. The package has capabilities for data transformations, variogram analyses, cross-validation, and univariate statistics (Table 3-1). Graphics in the package are limited to simple line-printer plots of the sample variogram points and data maps. The menu-driven package includes a tutorial using all of the kriging routines. The package is distributed with not all, but most source codes and, therefore, can be modified by the user if desired. All two-dimensional kriging routines can be executed from the command line, which provides users with the opportunity for batch processing.

c. The geostatistical environmental assessment software known as GEO-EAS (Englund and Sparks 1991) also is an interactive, menu-driven kriging software package for performing two-dimensional kriging. It has no direct provisions for universal kriging (Table 3-1). GEO-EAS does have an advantage over STATPAC in its enhanced graphics capabilities, which are useful in the interactive fitting of theoretical variograms to sample variogram points. In addition, in the computation of the sample variogram points, GEO-EAS allows for variable bin sizes, the use of which will be further discussed in Chapter 4.

d. STATPAC and GEO-EAS were originally developed for the personal computer. Since then, versions of GEO-EAS have been developed for some types of work stations. The kriging routines in STATPAC have not been adapted to work stations.

e. A third software package, the geostatistical software library known as GSLIB (Deutsch and Journel 1992), is a suite of programs developed over the years at Stanford University, Stanford, CA. It is presented as a collection of routines that are machine-independent (Table 3-1) and are intended to be used as a modular concept. The package is distributed as a suite of FORTRAN source codes that need to be compiled. Use of GSLIB requires a relatively high level of familiarity with geostatistics for its efficient use. As in the previous two software packages, GSLIB handles variogram analysis and kriging techniques (Table 3-1). Two of its primary advantages over the other two packages are its simulation techniques and ability to analyze three-dimensional data sets. Such techniques are useful especially in estimating potential extreme outcomes in a geostatistical analysis.

f. The Department of Defense Groundwater Modeling System (GMS) is a fourth software package that has kriging capabilities. GMS is a windows-based integrated modeling environment for site characterization, groundwater flow and transport modeling, and visualization of results. The GSLIB software has been implemented within

GMS to facilitate two- and three-dimensional kriging and interactive variogram modeling. GMS also provides comprehensive visualization techniques as well as other interpolation techniques that can be used as alternatives to kriging. The GMS system was developed for the Department of Defense by the Brigham Young University Engineering Computer Graphics Laboratory. GMS may be obtained from the U.S. Army Groundwater Modeling Technical Support Center, (U.S. Army Engineer Waterways Experiment Station, Vicksburg MS 39180).

g. A final note concerning geostatistical software and literature is that there can be differences in jargon or notation. These differences may cause some initial confusion if users or readers do not pay careful attention to the jargon or notation. For example, some authors may wish to use the term "semi-variogram" rather than "variogram"; others may express random variables as other than Z as has been done in this ETL, and it is common for different software to have different references for directional angles when discussing anisotropy.

Table 3-1
Geostatistical Software Characteristics

Characteristic	STATPAC	GEO-EAS	GSLIB	GMS2.0
Operating system	DOS	DOX/UNIX	Independent (requires FORTRAN compiler)	WINDOWS 95 UNIX
Menu-driven	Yes	Yes	No	Yes
Batch processing	Yes	No	Yes	Yes
User modifications	Yes, source code provided	No	Yes, source code provided	No
Data-set constraints	Yes, modifications possible via source code	Yes	Yes, modification possible via source code	Yes
ASCII output	Yes	Yes	Yes	Yes
Univariate statistics	Yes	Yes	Yes	Yes
Additional exploratory capabilities	Yes	Yes	Yes	Yes
Graphical support for analysis	Yes	Yes	Yes	Yes
Transformation	Yes	Yes	Yes	Yes
Back-transformation	No	No	Yes	Yes
Variogram construction	Yes	Yes	Yes	Yes
Variogram analysis	Yes	Yes	Yes	Yes
Variogram graphics	Yes	Yes	Yes	Yes
Cross-validation operations	Yes	Yes	Yes	Yes
Ordinary kriging	Yes	Yes	Yes	Yes
Universal kriging	Yes	No	Yes	Yes
Block kriging	Yes	Yes	Yes	Yes
Indicator kriging	Yes	Yes	Yes	No
Conditional simulation	Perhaps with batch processing	No	Yes	No
Three-dimensional kriging	Perhaps with batch processing	No	Yes	Yes
Mapping	Yes	Yes	Yes	Yes
Contouring	Yes	Yes	Yes	Yes
Gray-scale maps	Yes	Yes	Yes	Yes
Line printer	Yes	No	Yes	Yes
High-resolution screen	No	Yes	Yes via postscript	Yes
High-resolution printer	No	Yes	Yes	Yes
Postscript	No	No	Yes	Yes

Chapter 4

Practical Aspects of Variogram Construction and Interpretation

4-1. General

a. Chapter 2 presented the mathematical foundation for geostatistics and the kriging technique. One theme that pervades the technique is the importance of the theoretical variogram. The theoretical variogram, or what we will often refer to simply as the variogram, is a mathematical function or model which is fitted to sample-variogram points obtained from data. Permissible models, which include those given in Chapter 2, belong to a family of smooth curves having particular mathematical properties and are each specified by a set of parameters. Chapter 4 will describe a sequence of stages for estimating and investigating sample variogram points and a calibration procedure for specifying the parameters of the variogram model eventually fitted to the sample points. Although the calibration procedure is largely an objective means for evaluating theoretical variograms, the process of obtaining sample variogram points and finalizing a theoretical variogram remains an art as much as a science. An understanding of the material presented in Chapter 2 as well as professional judgment achieved through experience in geostatistical studies is important in effectively using the guidelines presented in this section.

b. An accurate estimate of a variogram is needed from a kriging perspective because the correlation matrix used to obtain the kriging weights is constructed from the variogram values. Even more directly, the variogram affects the computation of the kriging variance (Equations 2-36 and 2-47) through the product of the kriging weights and variogram values. An accurate variogram also has utility outside the strict context of kriging. For example, in augmenting a spatial network with new data collection sites, the range parameter of the variogram could be used as the minimum distance of separation between the new sites and between new and existing sites to maximize overall additional regional information. In another non-kriging-specific application, the variogram is used in dispersion variance computations in which the variance of areal or block values is estimated from the variance of point-data values (e.g., Isaaks and Srivastava (1989), p. 480).

c. The stages of variogram construction are described using an example data set of groundwater elevations measured near Saratoga, WY (Lenfest 1986), that are summarized in Table 4-1 and whose relative locations are shown in Figure 4-1.

d. The sequence of steps in computing sample variogram points depends on the stationarity properties of the regional variable represented by the data. If the mean of the regional variable is the same for all locations, then it is said to be spatially

Table 4-1
Univariate Statistics for Example Data Sets¹

Example Identifier	Number of Measurements	Transformation	Minimum (Base units)	Maximum (Base units)	Mean (Base units)	Median (Base units)	Standard Deviation (Base units)	Skewness (Dimensionless)
Saratoga		Drift	2,016.6	2,254.3	2,119.25	2,104.35	56.79	0.45
Water level A	83	Drift	25.6	65.68	42.30	38.54	10.13	1.03
Water level B	74	Drift	25.6	65.68	42.85	38.71	10.59	0.87
Bedrock A	108	None	22.64	80.48	44.42	42.82	10.76	0.89
Bedrock B	89	None	24.53	69.22	43.67	43.17	8.58	0.26
Water quality A	66	Natural log	2.08	8.01	5.19	5.59	1.75	-0.42

¹Base unit for Saratoga, water levels, A and B, and Bedrock A and B is feet; base unit for water quality A is log concentration, concentration in micrograms per liter.

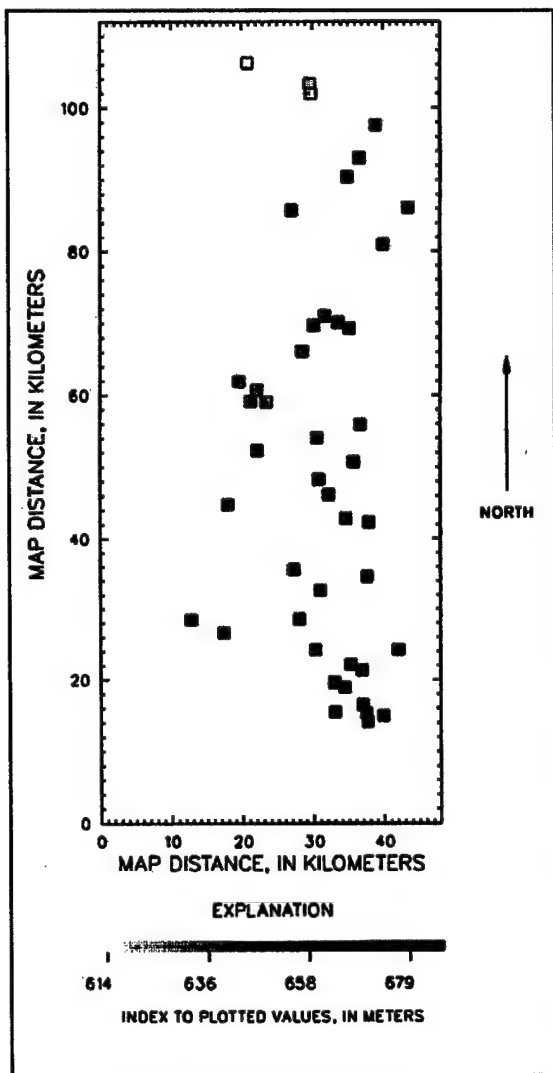


Figure 4-1. Measured water levels from Saratoga data

stationary; if the mean changes with location, then it is spatially nonstationary. Generally, if the data have a stationary spatial mean, the discussions in sections 4-3 and 4-7, which address nonstationarity and additional trend considerations, can be omitted. If the spatial mean is not stationary, as for this example data set, then sections 4-3 and 4-7 become important, and the sequence of stages for obtaining a variogram becomes an iterative procedure. All variogram and kriging computations for the Saratoga groundwater levels example were

performed by the interactive kriging software described in Grundy and Miesch (1987).

4-2. General Computation of Empirical Variogram

a. As described in section 2-3, the variogram $\gamma(h)$ characterizes the spatial continuity of a regional variable for pairs of locations as a function of distance or lag h between the locations. This variogram is sometimes called the theoretical variogram because it is assigned a continuous functional form that expresses the spatial correlation for any lag in the region of analysis. The function is estimated by fitting one of the equations given in section 2-3 to empirical or sample variogram points $\hat{\gamma}(h)$ using data whose locations contribute only a finite number of lags. Although $\hat{\gamma}(h)$ characterizes the spatial correlation of the data, it is computed from residuals of the data off the spatial mean. Therefore, without prior knowledge of nonstationarity in the underlying spatial process, the first step in computing the sample variogram is to identify existing nonstationarity indicated for the spatial mean.

b. The approximation to Equation 2-19 begins by computing squared differences $D_{i,j}^2$ from the data values $z(x_1), z(x_2), \dots, z(x_n)$ collected at locations x_1, x_2, \dots, x_n

$$D_{i,j}^2 = [z(x_i) - z(x_j)]^2 \quad (4-1)$$

If the spatial mean is stationary, then the squared differences of the data are equivalent to the squared differences of the residuals, and sample variogram computations can be continued using the data themselves. If the spatial mean is strongly nonstationary, the plot of Equation 4-1 versus the distance between associated points may indicate a trend or drift that would need to be removed before further variogram computations could be made. Drift would have to be considered in HTRW studies, such as determining contaminant concentrations areally dispersed from localized sources or

determining groundwater elevations following a local or regional gradient. In such studies sample variogram computations need to be made using residuals obtained by subtracting the estimated drift value at each location from the value of the datum at the location.

c. The data in Equation 4-1 are differenced without considering the relative direction between the locations; that is, D_{ij}^2 is isotropically computed. A plot of D_{ij}^2 versus h_{ij} for all i, j ($i > j$), where $h_{ij} = |x_i - x_j|$, produces a cloud of points whose properties govern the behavior of $\hat{\gamma}$. The central tendency of the cloud would generally increase with h . A substantial increase in the central tendency that persists for large h can indicate a nonstationary spatial mean. The cloud computed for the Saratoga data, with groundwater levels (z) in meters and distance (h) in kilometers, is shown in Figure 4-2 and does show increasing D^2 with increasing h , indicating potential non-stationarity.

d. Generally, there is a large amount of scatter in these plots, as seen in Figure 4-2, and this scatter can conceal the central behavior of D^2 with h . One way to estimate the central tendency and to minimize the effect of aberrant data values is to collect the D^2 into K bins or lag intervals of width $(\Delta h)_k$, $k=1, \dots, K$ and assign to $\hat{\gamma}$ the average of the values of D^2 in each bin. This process is similar to the way data are placed in bins for obtaining histograms. The expression for the k th average bin value is

$$\hat{\gamma}(h_k) = \frac{1}{2N(h_k)} \sum D_{ij}^2 I_k(h_{ij}) \quad (4-2)$$

where $N(h_k)$ is the number of squared differences that fall into bin k , and h_k is the lag distance associated with bin k . $I_k(h_{ij})$ is an "indicator function" that has a value of one if the h_{ij} falls into bin k and zero otherwise (it only includes values of D_{ij}^2 in the calculation that have an h_{ij} that falls into the bin). The lag value h_k can be the midpoint of the bin or it can be the average of the actual lag values for the points that fall in the bin.

e. To establish bins, either equal bin widths are specified and the distance between the two most separated data points, h_{max} , is subdivided according to these equal increments, or a K is chosen that defines the bin width. For the Saratoga data, a bin width of about 8 km established $K=12$ bins for γ . The $\hat{\gamma}$ points computed from the binned D_{ij}^2 values of Figure 4-2 are shown in Figure 4-3. The lag plotting positions are the average h values in the bin. The symbol x indicates that $N(h)$ is less than 30 pairs for the particular bin and this differentiation will be discussed in section 4-3. Although the sample variogram is still preliminary, its general behavior at this stage is adequate to indicate if nonstationarity needs to be addressed before sample variogram refinement is undertaken.

4-3. Nonstationarity

a. An indication of substantial nonstationarity or drift in the spatial mean would be a parabolic shape through all lags in a plot of $\hat{\gamma}$. This shape occurs because differences between data contain differences in the drift component that increase as h increases. If Equation 2-16 is inserted into Equation 2-17, squaring the differences in μ greatly amplifies the increase with h . In these cases of drift, generally a low-order (less than three) polynomial drift in (u, v) is fitted to the data and subsequently subtracted from the data to obtain residuals. Trend surfaces are not necessarily limited to polynomial forms. For example, a numerical model of groundwater flow may be used to obtain residuals of groundwater head data.

b. In theory, the polynomial trend reflects a slowly varying drift in the spatial mean and, as such, one regional trend surface should be fitted to all the data. However, often the drift and residuals are obtained locally; that is, using moving neighborhoods of locations. Estimates of these values at any point are thus made using a reduced number (usually between 8 and 16) of surrounding locations. This is done because ultimately the kriging estimates are made using only the data values in the given neighborhood. Manipulating the kriging

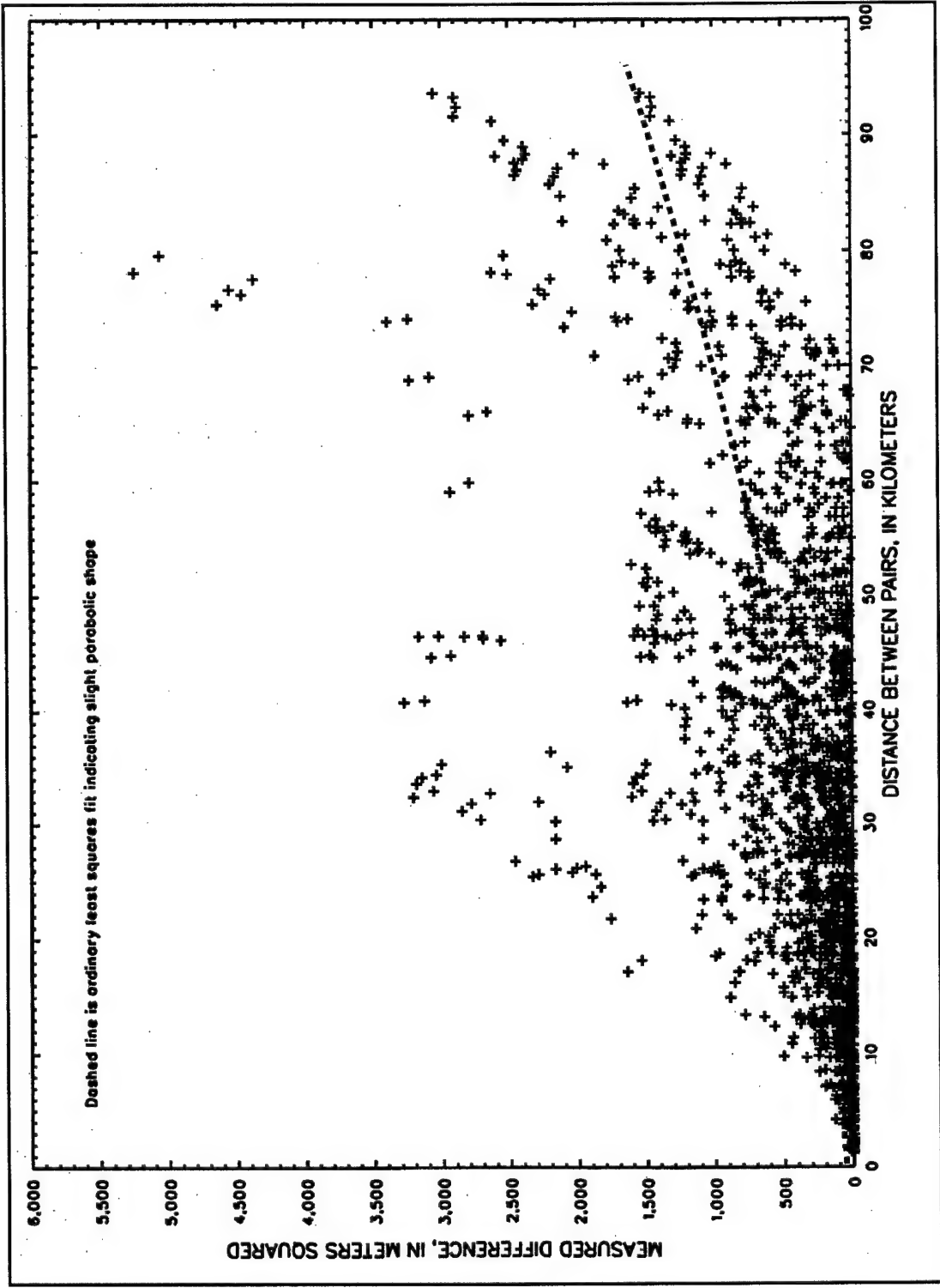


Figure 4-2. Squared differencing of values for all possible pairs of points for Saratoga data

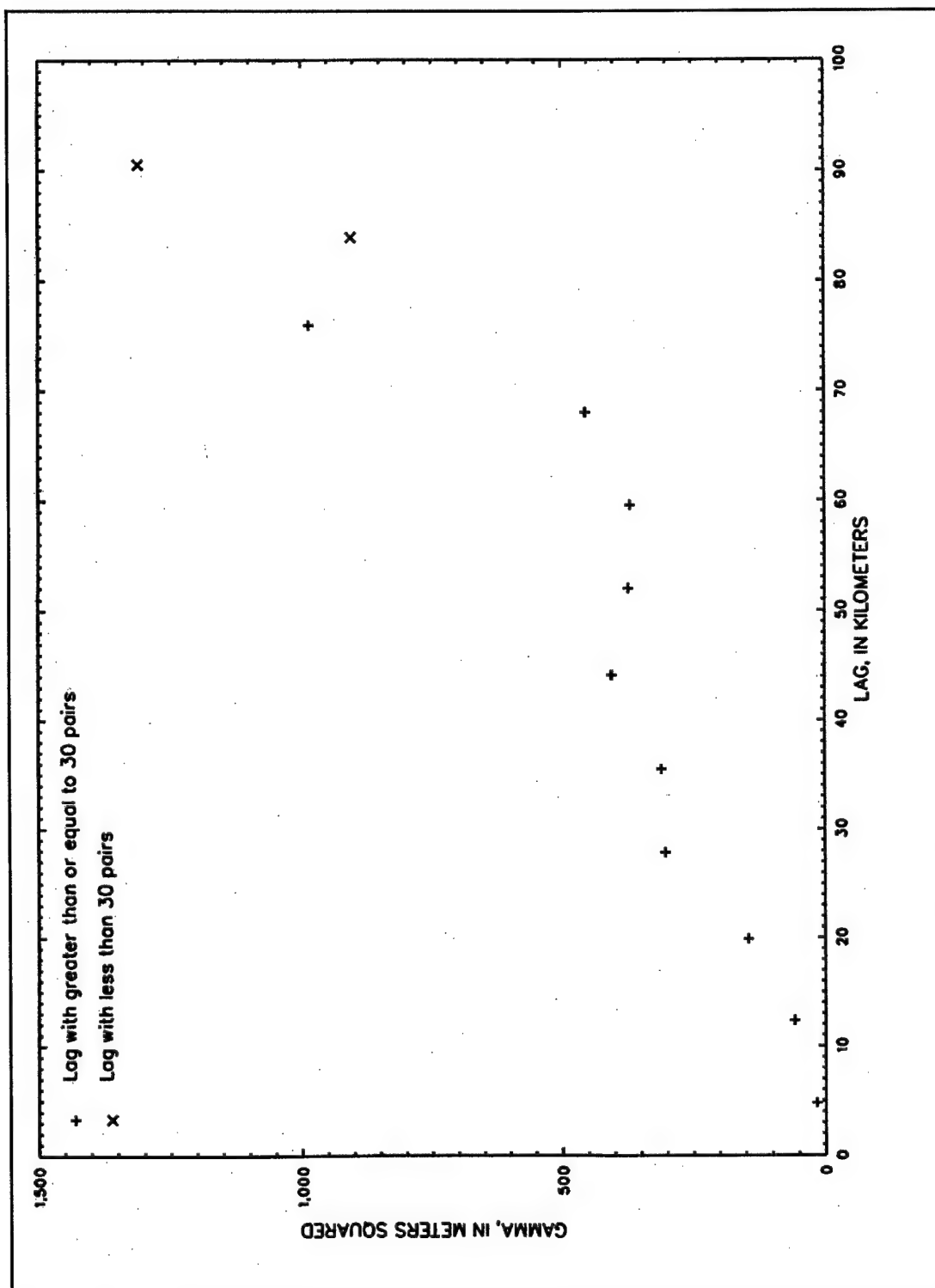


Figure 4-3. Initial sample variogram points for Saratoga data

matrices takes less time when a smaller number of data values are used to make estimates and these efficiencies can be significant when dealing with large data sets. Little accuracy is lost because the nearest neighbors are the most influential in the kriging weighting scheme.

c. A parabolic shape to $\hat{\gamma}$ for the Saratoga data is shown in Figure 4-3 for the sample variogram points plotted for lags up to about 32 km (the first four points) and for lags beyond about 56 km. The presence of a parabolic shape in the sample variogram points was not surprising, because examination of the data indicates a north-south gradient in the groundwater levels. The simplest polynomial trend, linear in u and v , was fitted to all the data using ordinary least-squares estimation. Residuals obtained by subtracting this regional trend surface from the data were used to reestimate $\hat{\gamma}$ in Equation 4-2 and the sample variogram for the residuals is shown in Figure 4-4.

4-4. Variogram Refinement

a. In the previous section, an initial $\hat{\gamma}$ was specified by points computed from Equation 4-2. In general, the larger $N(h_k)$ is for any bin or lag interval k , the more reliable will be the points defining $\hat{\gamma}(h_k)$. Also, the larger K is, the greater the number of sample variogram points shaping $\hat{\gamma}$. However, $N(h_k)$ and K are competing elements of $\hat{\gamma}$. Journel and Huijbregts (1978) suggest that each lag interval k should have $N(h_k)$ equal to at least 30 pairs. The American Society for Testing and Materials (Standard D5922-96) suggests 20 pairs for each lag interval. For small data sets the number of intervals may have to be small to guarantee either number of recommended pairs in all intervals.

b. It is difficult to determine the minimum number of data values n needed to satisfy the $N(h_k)$ requirements for all lag intervals of a sample variogram. Simple combinatorial analysis can establish a sample size needed to achieve a given total number of distinct pairs of items taken from the sample, but it does not address the spatial

considerations needed for proper lagging. As an example, for data collected on a uniform grid and equal-sized bins, fixing an n to just satisfy the minimum $N(h_k)$ for the smaller lags will yield insufficient data pairs to meet the minimum $N(h_k)$ for the larger lags. Fixing an n to assure the minimum $N(h_k)$ for the larger lags will generally have $N(h_k)$ much greater than the minimum for the smaller lags. Therefore, the question of how much data is required to adequately compute a variogram should also address the relative locations of the data-collection sites.

c. The first 10 of the 12 bins for $\hat{\gamma}$ for the Saratoga data contained more than 30 data pairs. Therefore, the bin width can be decreased to get more points defining the early part of $\hat{\gamma}$. These bin-width adjustments can be made to refine $\hat{\gamma}$ whether it is computed from the data or from the residuals. A plot of $\hat{\gamma}$ for the residuals for the Saratoga groundwater elevations with the bin width narrowed to about 6.5 km is shown in Figure 4-5.

d. Spatial data are usually not collected on a uniform grid but occur in a pattern that reflects problem areas, accessibility, and general spatial coverage. In the Saratoga data set, nonuniform data spacing results in the number of data pairs in each bin, although still greater than 30, being highly variable among the bins. This variability yields different reliabilities for the points defining $\hat{\gamma}$. To establish a balance for $N(h_k)$ among the bins, variable bin sizes can be used so that each bin contains approximately the same (large) number of points. A bin with fewer points can be coalesced with an adjacent bin to form a wider bin with a greater number of points. Conversely, a bin with an excessive number of points can be subdivided into adjacent, narrower bins. The coalescing and subdividing procedure is largely trial and error, until the distribution of the pairs of points is satisfactory to the investigator.

e. The values of $\hat{\gamma}$ at the smaller lag values are the most critical to define the appropriate γ . Therefore, the trade-off between the number of bins and the number of data pairs within each bin can be varied for different regions of the sample

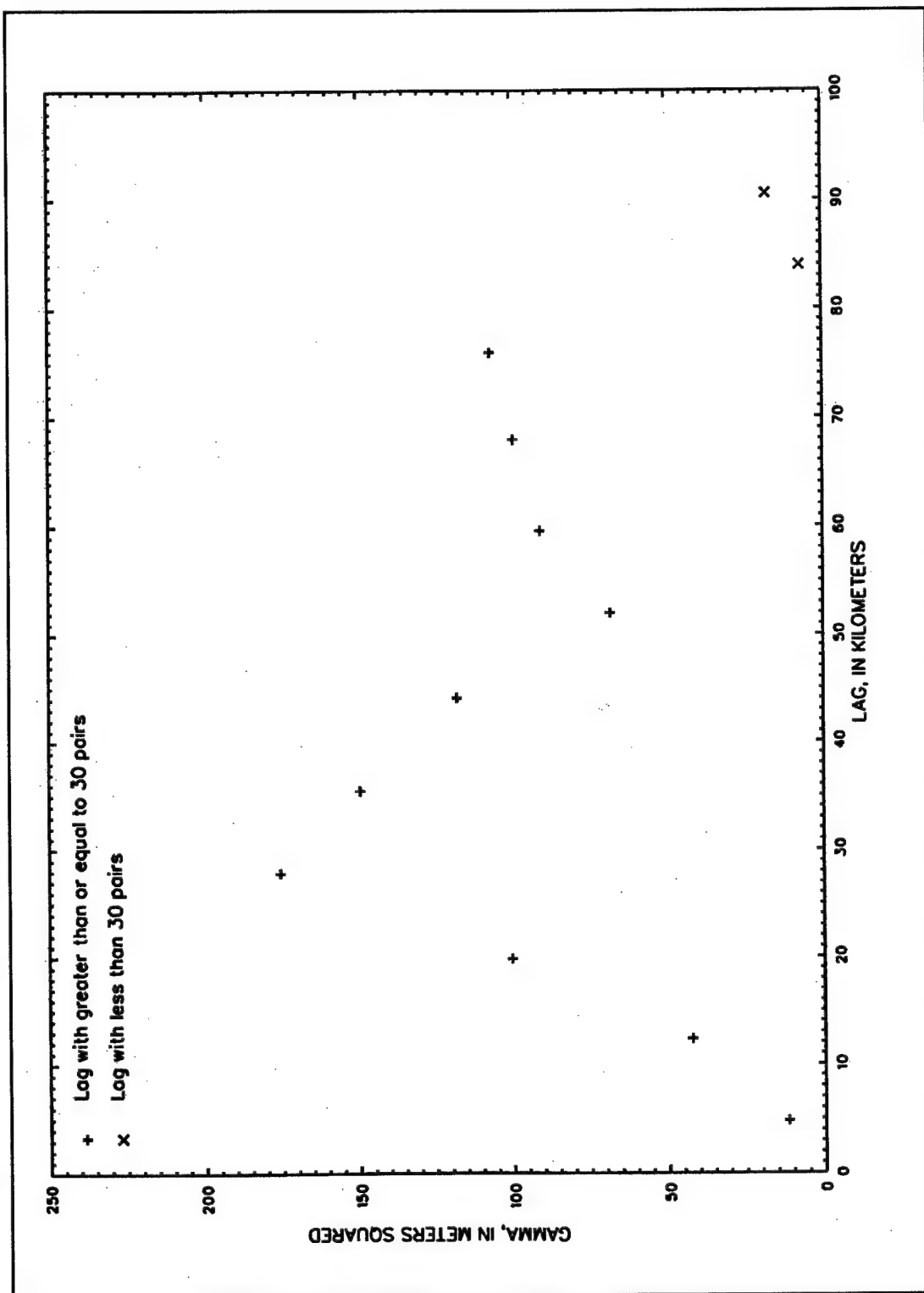


Figure 4-4. Sample variogram points for ordinary-least squares trend residuals of Saratoga data

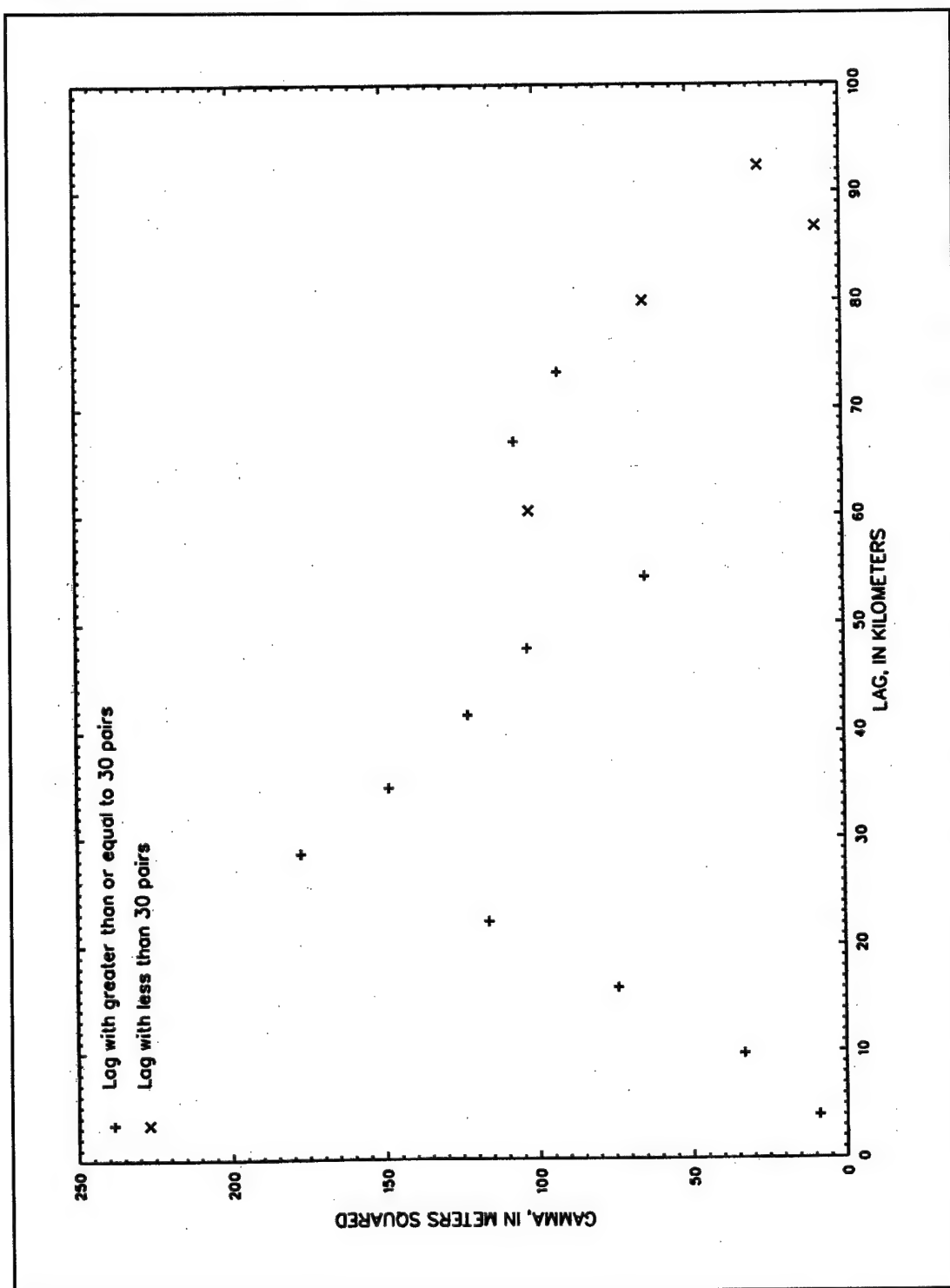


Figure 4-5. Sample variogram points for ordinary-least squares trend residuals of Saratoga data binned to 4 miles

variogram. At smaller lags, the numbers of data pairs per bin can be nearer the minimum $N(h_k)$ to define more bins. At larger lags, a smaller number of wider bins would be adequate. Knowing that the variogram should be a smooth function, ultimately the analyst visually decides when the sample variogram is sufficiently defined at all lags to adequately approximate a theoretical variogram.

4-5. Transformations and Anisotropy Considerations

a. Transformations.

(1) A transformation is applied to a data set generally for one of two interrelated purposes. First, a transformation can reduce the scale of variability of highly fluctuating data. This variability would especially occur with contaminant concentrations in which order of magnitude changes in data at proximate sites are not uncommon. The effects of such data would be erratic sample variogram points as exhibited by a large-amplitude, ill-defined sawtooth pattern of the lines connecting the points.

(2) Second, a proper transformation of data whose probability distribution is highly skewed often produces a set of values that is approximately normally distributed by mitigating the influence of problematic extreme data values. A data set with a normal distribution is important in kriging when confidence levels of the estimates are desired. This usage of confidence levels in a kriging analysis will be illustrated in Chapter 5.

(3) Among the more common transformations is the natural log transform. As an example, for this transformation, the $\hat{\gamma}$ will be the sample variogram values of logarithms, and subsequent kriged estimates will be logarithms. Another transformation that is often used, especially in spatial analyses of contaminant levels, is the indicator transformation described in Chapter 2. Although a transformation might achieve better-behaved sample variogram points, there are subtleties to

consider in interpreting the kriging results of the transformed data or in back-transforming kriging results into the untransformed (original) units, as discussed in Chapter 1. If a satisfactory variogram of the original data cannot be achieved and a transformation is indicated, the sample variogram computation process must begin again with Equation 4-2. Even though no transformation was needed for the Saratoga data, an example using a logarithmic transformation and an example using the indicator transformation are presented in Chapter 5.

b. Directional variograms and anisotropy.

(1) Anisotropy in the data can be investigated by computing sample variograms for specific directions. Locations included in a given direction from any other location are contained in a sector of a circle of radius h_{max} centered on the location. The sector is specified by two angular inputs. The first is a bearing defining the specific direction of interest [measured counterclockwise from east ($=0^\circ$)] and the second is a (window) angle defining an arc of rotation swept in both directions from the bearing. Thus, in the terminology used here, the total angle defining a direction is equal to twice the window angle. Differences in sample variograms computed using these angle windows specified for different directions can be an indication of anisotropy.

(2) Anisotropy is generally either geometric or zonal. Geometric anisotropy is indicated by directional theoretical variograms that have a common sill value, but different ranges. The treatment of geometric anisotropy is dependent on the software used. The lags of the directional variograms can be scaled by the ratio of their ranges to the range of a standard or common variogram. In some cases, the lags of all directional variograms are scaled by their respective ranges, and a common variogram with a range of 1 is used. Groundwater contaminant plumes often have geometric anisotropy in which the prevailing plume direction would have a greater range than that of the transect of the plume.

(3) Zonal anisotropy is indicated by directional variograms that have the same range but different sills. Pure zonal anisotropy is usually not seen in practice; typically it is found in combination with geometric anisotropy. Such mixed anisotropy may be encountered if evaluating the variograms of three-dimensional HTRW sampling results. Variability of such data (as indicated by the sill of the variogram) may be significantly higher and the range significantly shorter in the vertical direction than in the horizontal direction. In order to model this mixture of anisotropic variograms, the overall variogram is set to a weighted sum of individual models of the directional variograms scaled by their ranges. In this process, called nesting, the choice of weights requires a trial and error approach with a constraint that the sum of the weights equals the sill of the overall variogram. The reader is referred to Isaaks and Srivastava (1989, pp. 377-390) for further information on both types of anisotropy.

(4) For a given number of data locations, directional sample variograms will necessarily have fewer points for any lag when compared to the points for the same lag in the omnidirectional variogram. Hence, there will be less reliability in the directional-variogram point values, which would be a critical constraining factor for small data sets or for a data pattern that does not conform to a direction of anisotropy. For a general idea of the sufficiency of the data to adequately determine any anisotropy, the computations of anisotropic sample variograms can be initially limited to two orthogonal directions with window angles of 45 deg.

(5) Directional sample variograms also can be used to further delineate nonstationarity of the spatial mean. If the omnidirectional sample variogram indicates a drift in the data, the directional variograms may determine the dimensionality of the drift. That is, although they may not establish the degree of the polynomial in the drift equation, the directional sample variograms can indicate the relative strengths of the drift in the u and v directions.

(6) The computed sample variograms for the general north-south and east-west directions for the Saratoga data are shown in Figure 4-6. The north-south variogram is specified by a direction angle of 90 deg and a window angle of 45 deg. The north-south variogram reveals the preferential north-south data alignment by mimicking the omnidirectional (direction angle = 0 deg and window angle = 90 deg) sample variogram of Figure 4-3. The east-west variogram is specified by a direction angle of 0 deg and a window angle of 45 deg. The lack of pairs of locations for the east-west variogram precludes a good analysis for this direction, but the overlap of the few sufficiently defined variogram points with the north-south variogram indicates a consistency of drift in the two directions. Because of this consistency, an isotropic variogram is assumed for the Saratoga residuals. An example of anisotropic variograms is described in Chapter 5.

4-6. Fitting a Theoretical Variogram to the Sample Variogram Points

a. General.

(1) The importance of adequately defining the bin values of a sample variogram is substantiated by the need to accurately generalize the data-based behavior of the sample variogram by a theoretical variogram γ . The parameters controlling the specific behavior of theoretical variograms are the nugget value, the range, the sill, or in the case of a linear variogram, a slope parameter. Of these parameters, the nugget and the sill can be related to properties and statistics of the data.

(2) The nugget is essentially the extrapolation of the sample variogram to a lag of zero. It reflects the uncertainty of the variogram at lags that are much smaller than the minimum separation between any two data locations. The nugget value can include measurement error variance, and an estimate of this variance will approximate a minimum value of the extrapolation.

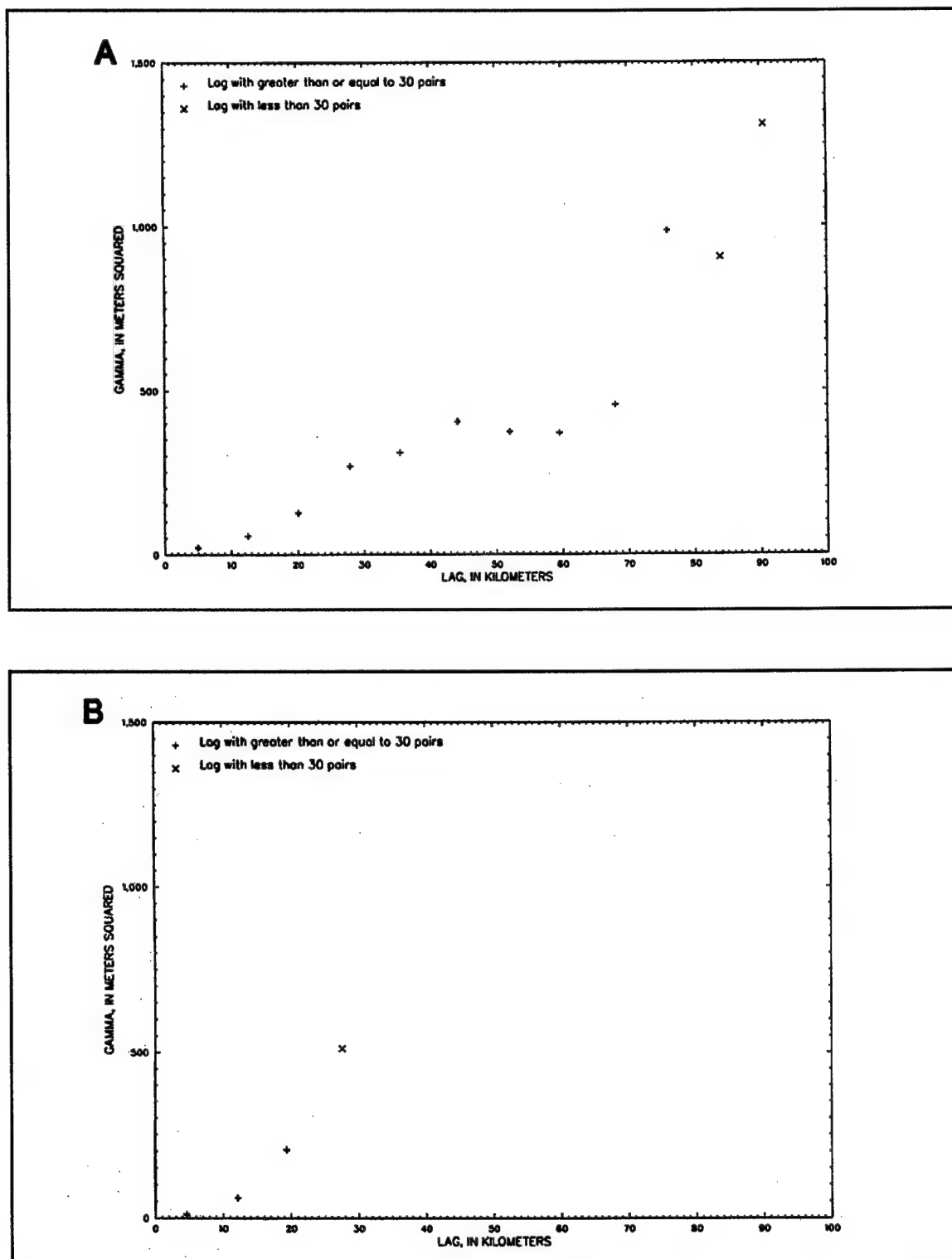


Figure 4-6. Initial directional sample variogram points for raw Saratoga data--A, north-south and B, east-west

(3) The sill determines the maximum value of a variogram and approximates the variance of the data. However, the points defining $\hat{\gamma}$ take precedence over the sample variance in locating the sill. Some variograms are unbounded, and others may only reach a sill value asymptotically. A defined sill allows conversion of the variogram to a covariance function using Equation 2-27, which is generally done because computations in the kriging algorithms are more efficiently performed using a covariance function.

(4) Fitting a function to the sample variogram values can range from a visual fit to a sophisticated statistical fit. A statistical fit is an objective method as long as the choice of bins and weighting of the sample variogram points remain fixed. However, because the inputs will vary with investigators, inherent subjectivity persists as in a visual fit. A final calibration of the variogram parameters would be based on the kriging algorithm and, thus, either of the initial fitting methods at this stage would suffice.

(5) Because the initial part of the variogram has the most effect on subsequent kriging output, a good estimate of the nugget value becomes a most important first step. The range and the sill, in that order, complete the ranking of the influence of variogram parameters on the output of a geostatistical analysis. Whatever the fitting method used, the theoretical variogram needs to be supported by the sample variogram values. For variograms with a range parameter, this support should extend to the range. Journel and Huijbregts (1978) suggest that this support should be through one-half the dimension of the field or essentially through one-half the maximum lag distance of the sample data.

(6) Most geostatistical studies can be successfully completed using the following four singular theoretical variogram forms: exponential, spherical, Gaussian, and linear functions (Figure 2-3). For the example variogram determination described in this section, only one of these singular forms will be selected; however, positive linear combinations of these forms also are acceptable as theoretical variograms (see section 4-5b).

Geometric relationships to aid in obtaining parameters for the four variogram forms are described in the following sections and are illustrated in Figure 2-3 for reference.

b. Exponential variogram.

The exponential variogram (Equation 2-23) is specified by the nugget g , sill s , and a practical range value r . The range is qualified as practical because the sill is reached only asymptotically. The initial behavior of the exponential variogram is different from the behavior of the spherical variogram in that the convex behavior extends to the nugget value (Figure 2-3). Again, a nugget value and a sill value are first specified based on the $\hat{\gamma}$ points. The practical range is chosen so that the value of the resulting exponential function evaluated at the practical range lag is 95 percent of the sill value. The specified exponential function would mesh with the sample variogram points at least through the practical range lag. An initial estimate of the practical range can be made by checking if the intersection of the sill value with a line tangent to the variogram at the nugget is at a lag value equal to one-third of the assumed practical range value as illustrated in Figure 2-3. Examples of the exponential variogram may be found in spatial studies of sulfate and total alkalinity in groundwater systems (Myers et al. 1980).

c. Spherical variogram. The spherical variogram parameters (Equation 2-24) are a nugget value g , a range r , and a sill s . At smaller lag values the sample variogram points indicate linear behavior from the nugget that then becomes convex and reaches a sill value at some finite lag (Figure 2-3). A sill is estimated, and a line drawn through the points of the initial linear part of the variogram would intersect the sill at a lag value approximately equal to two-thirds of the range. With these estimates of the parameters, a spherical variogram is defined that should be supported by the sample variogram points. If the spherical plot does not fall near the sample variogram points, adjustments need to be made to the parameter estimates and the subsequent fit evaluated. Although the spherical variogram is one of the most often

used models for real valued spatial studies, it seems to be a predominant model for indicator values at various cutoff levels as, for example, in a study of lead contamination (Journel 1993).

d. Gaussian variogram. The Gaussian variogram parameters (Equation 2-25) are a nugget value g , and a sill s , and this variogram also has a practical range r . The Gaussian variogram is horizontal from the nugget, becomes a concave upward function at small lags, inflects to concave downward, and asymptotically approaches a sill value (Figure 2-3). After a nugget value and sill value are specified based on the points, the variogram value at a lag of one-half the estimated practical range will be two-thirds of the sill value. Again, this fitted variogram needs to be supported by the $\hat{\gamma}$ points to a reasonable degree. As will be described in the example using the Saratoga data, the Gaussian variogram often is used where the variable analyzed is spatially very continuous, such as a groundwater potentiometric surface.

e. Linear variogram. Parameters for a linear variogram (Equation 2-26) are a nugget value g , and a slope b . Sample points indicating a linear variogram would increase linearly from the nugget value and fail to reach a sill even for large lags (Figure 2-3). With the nugget as the intercept, the slope is computed for the line passing through the $\hat{\gamma}$ points. A pseudosill s can be defined as the value of the line at the greatest lag, h_{max} , between any two locations. This lag becomes the defacto range r for a linear variogram. Examples of the usage of the linear variogram occur in hydrogeochemical studies of specific conductance and in studies of trace elements such as barium and boron (Myers et al. 1980).

4-7. Additional Trend Considerations

a. If a drift in the data is indicated as in section 4-3, the theoretical variogram of residuals that has been fitted thus far is used to update the drift equation. Although ordinary least squares often suffices for computing a polynomial drift equation, drift determination itself is a function of

γ when the data are spatially correlated. But γ cannot be estimated until a drift equation is obtained to yield the residuals. Therefore, obtaining a sample variogram and a subsequent theoretical variogram from drift residuals of a specified drift form is an iterative process (David (1977), pp. 273-274) framed by the following steps:

(1) An initial variogram is specified and drift coefficients are computed to obtain residuals. For this step, a pure nugget (i.e. constant) variogram can be used to compute the initial estimates of the drift coefficients. This is an ordinary least-squares estimate of the drift yielding a first-iteration sample variogram of residuals.

(2) A theoretical variogram is fitted to the sample variogram of the residuals and is used to obtain updated drift coefficients.

(3) The residuals from the drift obtained in step b are used to compute an updated sample variogram.

(4) The sample variogram computed at the end of step 3 is compared to the sample variogram of step 2. If the two sample variograms compare favorably, then the theoretical variogram from step 2 is accepted as the variogram of residuals for subsequent kriging computations. If the sample variogram from step 3 differs markedly from the sample variogram of step 2, steps 2-4 are repeated using the sample variogram of the most recent step c.

b. Generally, the plot of the points of $\hat{\gamma}$ from a set of residuals will initially increase with h , reach a maximum, and then decrease as seen in Figure 4-4. This typical haystack-type behavior, discussed by David (1977, pp. 272-273), is attributed to a bias resulting from the estimation error in the drift form and its coefficients. Thus, this behavior in the variogram of the residuals generally would more readily occur with a higher degree of drift polynomial. This behavior should not prohibit acceptable variogram determination because the initial points of the sample variogram of residuals are still indicative of the theoretical

variogram. For example, the lag associated with the maximum of $\hat{\gamma}$ of the residuals can be a good first approximation for the range of the theoretical variogram.

4-8. Outlier Detection

a. Outliers in a data set can have a substantial adverse effect on $\hat{\gamma}$. However, divergent data values can be screened for evaluation using a Hawkins statistic (Hawkins 1980), which is described in the context of kriging by Krige and Magri (1982). A neighborhood containing 4 to 10 data points, approximately normally distributed, around each suspected outlier must be defined. Despite potential outliers in the data set, a best guess initial theoretical variogram also is needed.

b. The Hawkins statistic is obtained by comparing a suspect datum to the mean value of the 4 to 10 surrounding data, the smaller number being sufficient if the variability is lower. The spacing between these surrounding points is accounted for by the properties of the chosen variogram. A value for the statistic of 3.84 or higher would indicate an outlier on the basis of a 95-percent confidence interval. A larger number of surrounding points has the direct effect of increasing the magnitude of the statistic. Anomalous points are removed from the data set and the procedures described for obtaining the sample variogram are repeated for the smaller data set. There were no outlier problems in the Saratoga data.

c. There is debate among geostatisticians regarding the merit of automated outlier-detection methods. A procedure such as that described here is presented as an investigative tool with the understanding that the investigator will also use attendant justification along with a Hawkins-type statistic to ultimately decide if a data value is discarded as a true outlier or retained as a valid observation. In some situations, highly problematic data values are removed for computation of the sample variogram points but are reinstated for kriging.

4-9. Cross-Validation for Model Verification

a. General.

(1) Parameters of the theoretical variogram obtained from the initial fitting and refinement of the sample variogram are calibrated using a kriging cross-validation technique. In this procedure, the fitted theoretical variogram is used in a kriging analysis in which data values are individually suppressed and estimates made at the location using subsets of the remaining points. As described in section 4-3, these subsets are the data points in a moving neighborhood surrounding the point under consideration. The calibration estimate made at each data location requires a matrix inversion, which could be very time-consuming if all remaining data locations were used to construct the matrices rather than just those within a neighborhood of a limited search radius.

(2) After kriged values at all data locations have been estimated in the above manner, the data are used with their kriged values and kriging standard deviation to obtain cross-validation statistics. A successful calibration is based on criteria for these statistics, which are described in the next section. If the criteria cannot be reasonably met by adjusting the parameters in the given theoretical variogram function, then calibration should be reinitialized with a different theoretical variogram function. In some data sets with nonstationary spatial means, the drift polynomial may have to be changed as well as the variogram to achieve a satisfactory calibration.

b. Calibration statistics.

(1) The kriging cross-validation error e_i corresponding to measurement $z(x_i)$ is defined as

$$e_i = z(x_i) - \hat{z}(x_i) \quad (4-3)$$

where $\hat{z}(x_i)$ is the kriged estimate of $z(x_i)$ based on the remaining $n-1$ measurements in the data set.

The kriged estimate is obtained by ordinary kriging if the spatial mean is constant or by universal kriging if the spatial mean is not stationary. A reasonable criterion for selecting a theoretical variogram would be to minimize the squared errors, $\sum e_i^2$, with respect to the variogram parameters. However, unlike ordinary least-squares regression, which also minimizes the sum of squared errors, simply minimizing the squared errors is not sufficient for kriging because the resulting model can yield highly biased estimates of the kriging variances, $\sigma_k^2(\mathbf{x})$, where $\sigma_k^2(\mathbf{x})$ is the kriging variance at location \mathbf{x} . This simple minimization would give unrealistic measures of the accuracy of the kriging estimates. To guard against such bias, an expression for the square of a reduced kriging error is defined:

$$\tilde{e}_i^2 = \frac{e_i^2}{\sigma_k^2(\mathbf{x}_i)} \quad (4-4)$$

where the kriging variances are computed using either Equation 2-36 or 2-47. If the kriging variance is an unbiased estimate of the true mean-squared error of estimate, then the reduced kriging errors would have an average near one. Therefore, the standard cross-validation procedure for evaluating a theoretical variogram is:

$$\begin{aligned} \min \left(\frac{1}{n} \sum_{i=1}^n e_i^2 \right)^{0.5} \\ \text{subject to } \left(\frac{1}{n} \sum_{i=1}^n \tilde{e}_i^2 \right)^{0.5} \approx 1 \end{aligned} \quad (4-5)$$

(2) The expression to be minimized is called the kriging root-mean-squared error and the constraint is called the reduced root-mean-squared error. The reduced root-mean-squared error should be well within the interval having endpoints

$1 + \left(2\sqrt{\frac{2}{n}} \right)$ and $1 - \left(2\sqrt{\frac{2}{n}} \right)$ (Delhomme 1978). An additional check on the cross-validation

results is the unbiasedness condition where

$$\frac{1}{n} \sum e_i \approx 0.$$

(3) As indicated in Chapter 2, if probabilistic statements concerning an actual value of Z at an unmeasured location are to be made relative to the kriged estimate and the kriging variance at the location, it is necessary to explore the distribution of the cross-validation kriging errors. In particular, it is desirable that the reduced errors, \tilde{e}_i , $i=1,2,\dots,n$, are approximately normally distributed with mean 0 and variance 1. A histogram or normal probability plot of the reduced kriging errors can be used to assess the validity of assuming a standard normal distribution for the reduced kriging errors. Additionally, if the distribution of reduced kriging errors can be assumed to be standard normal, outliers not detected using the method discussed in section 4-7 may be detected by comparing the absolute values of the reduced kriging errors to quantiles of the standard normal distribution.

(4) Using the Saratoga data, a spherical variogram was fitted to the refined sample variogram of the residuals. The estimated nugget was about 1.49 m², the sill was 133.8 m², and the range was about 48 km. Because of difficulty in determining an exact extrapolated value for the nugget, the value of 1.49 m² was selected based on an estimated measurement error related to obtaining water levels at the well depths in the Saratoga valley.

(5) After two iterations using drift residuals, as described in section 4-7, a final variogram was chosen with a nugget of 1.49 m², a sill of 148.6 m², and a range of 44.8 km (Figure 4-7). These parameters defined the theoretical variogram used to obtain the cross-validation errors using universal kriging with an assumed linear drift. The best combination of statistics that could be obtained after several attempts at refining the model were a root-mean-squared error of 3.45 m and a reduced root-mean-squared error of 0.5794. The

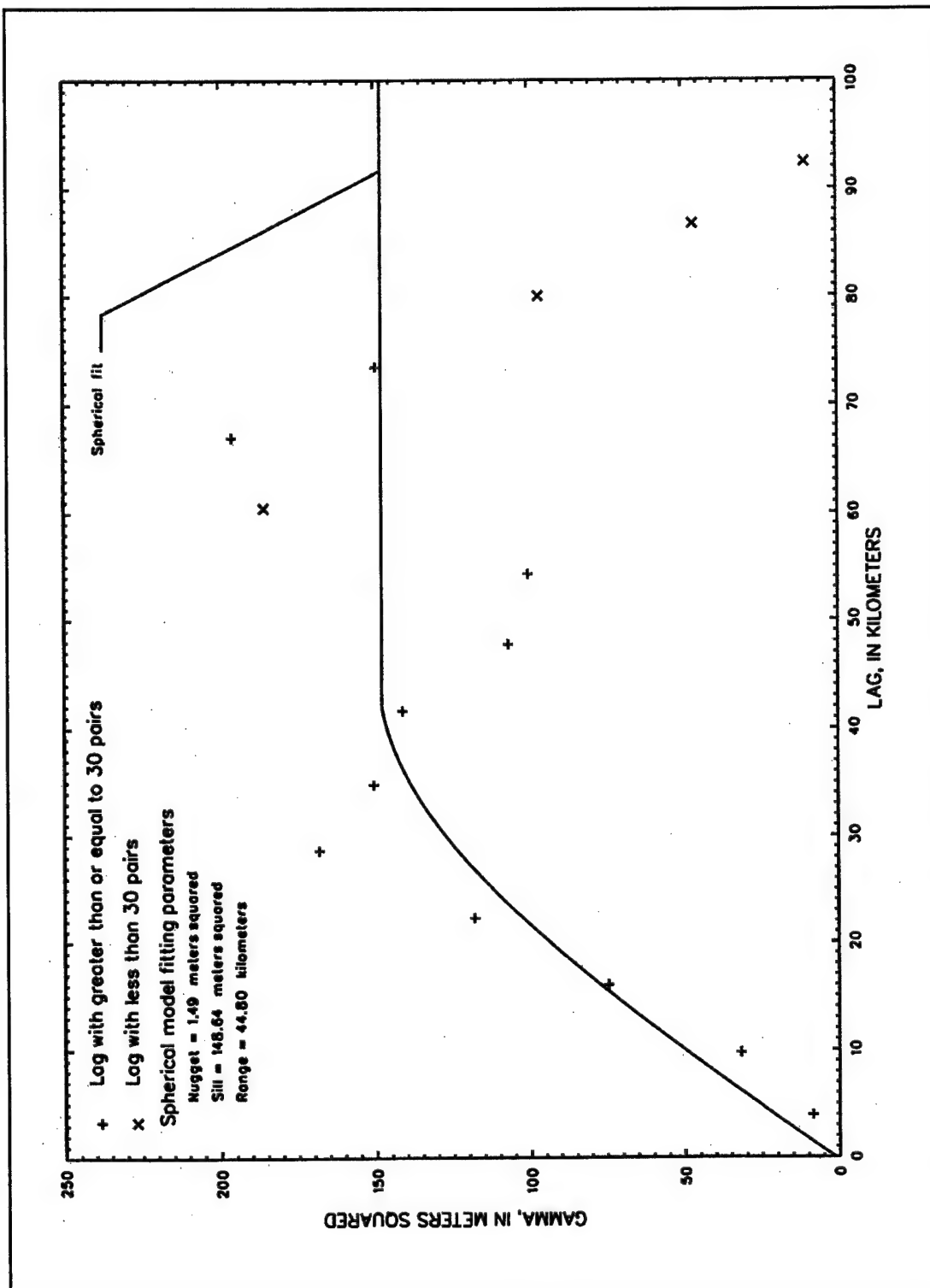


Figure 4-7. Sample variogram points and theoretical spherical fit for Iterated Saratoga residuals

reduced-root-mean-squared error is too small, indicating that the kriging variances produced by the model are too large compared to the actual squared errors. This fact, coupled with the rather large root-mean-squared error, makes the theoretical variogram model unacceptable. In section 4-9c, a Gaussian variogram is fitted to the data that produces much better cross-validation results than the results for the spherical variogram.

c. Variogram-parameter adjustments.

(1) If any of the cross-validation statistics vary unacceptably from their suggested values, minor adjustments to the variogram parameters can be made to attempt to improve the statistics. Whatever modifications are made to the parameters, they should not have to be so severe that the variogram function drastically deviates from the sample variogram points. If the support of the sample variogram points is compromised in order to achieve acceptable cross-validation results with the given drift-variogram model, a different drift-variogram combination should be investigated.

(2) A reduced root-mean-squared error that is unacceptable may be improved upon by adjusting the range parameter or the nugget value of the variogram. Modifying the range parameter should be considered first and any shifts in the nugget value should be minimal and made only as a final recourse. Calibration errors are relatively insensitive to minor adjustments of the sill.

(3) If the reduced root-mean-squared error is too small, as in the Saratoga example, extending the range (equivalent to decreasing the slope for a linear variogram) will decrease the kriging variance and thus increase the reduced root-mean-squared error. If a shift in the nugget value is required, a decrease in the nugget will reduce the kriging variance. If the reduced root-mean-squared error is too large, then a contraction of the range or a positive shift in the nugget value can be made, keeping in mind the above caveat of priority and extent of the changes. Changes in these

parameters, generally, also will have an effect on the mean-squared error. The larger the nugget is as a percentage of the sill, the larger the mean-squared error will be. In general, improvements in one statistic are usually made at the expense of the other statistics. The optimization of the statistics as a set is, in effect, a trial and error procedure that is operationally convergent.

(4) Reduced kriging errors may not approximate a standard normal distribution. If this is the case, a transformation of the data may be needed to achieve a more normal distribution, and the variogram estimation procedure would be repeated.

(5) Because no convergence could be reached for parameter values of a spherical variogram for the Saratoga data, a Gaussian theoretical variogram was fitted to the sample variogram of residuals in Figure 4-4. This choice was made because the initial sample variogram points could be interpreted to have a slight upward concavity, but eventually reached a sill. This behavior can be attributed to correlation rather than to further drift. After an iterated cross-validation with the Gaussian parameters, a Gaussian variogram with a nugget of 1.49 m^2 , a sill of 185.81 m^2 , and a range of 27.52 km (Figure 4-8) yielded a root-mean-squared error of 2.33 m and a reduced-root-mean-squared error of 1.083 . The mean cross-validation error is 0.0195 m . These values represent an improvement over the spherical variogram and were deemed acceptable for the Gaussian variogram.

(6) A probability plot of the reduced kriging errors using the final Gaussian variogram is shown in Figure 4-9. It is reasonably linear between two standard deviations and, thus, approximates a standard-normal-distribution function. Finally, a plot in Figure 4-10 of the data versus their kriged estimates indicates that the linear drift-Gaussian variogram model selected for the Saratoga data would produce accurate estimates of groundwater elevations for interpolation or contour gridding in the region.

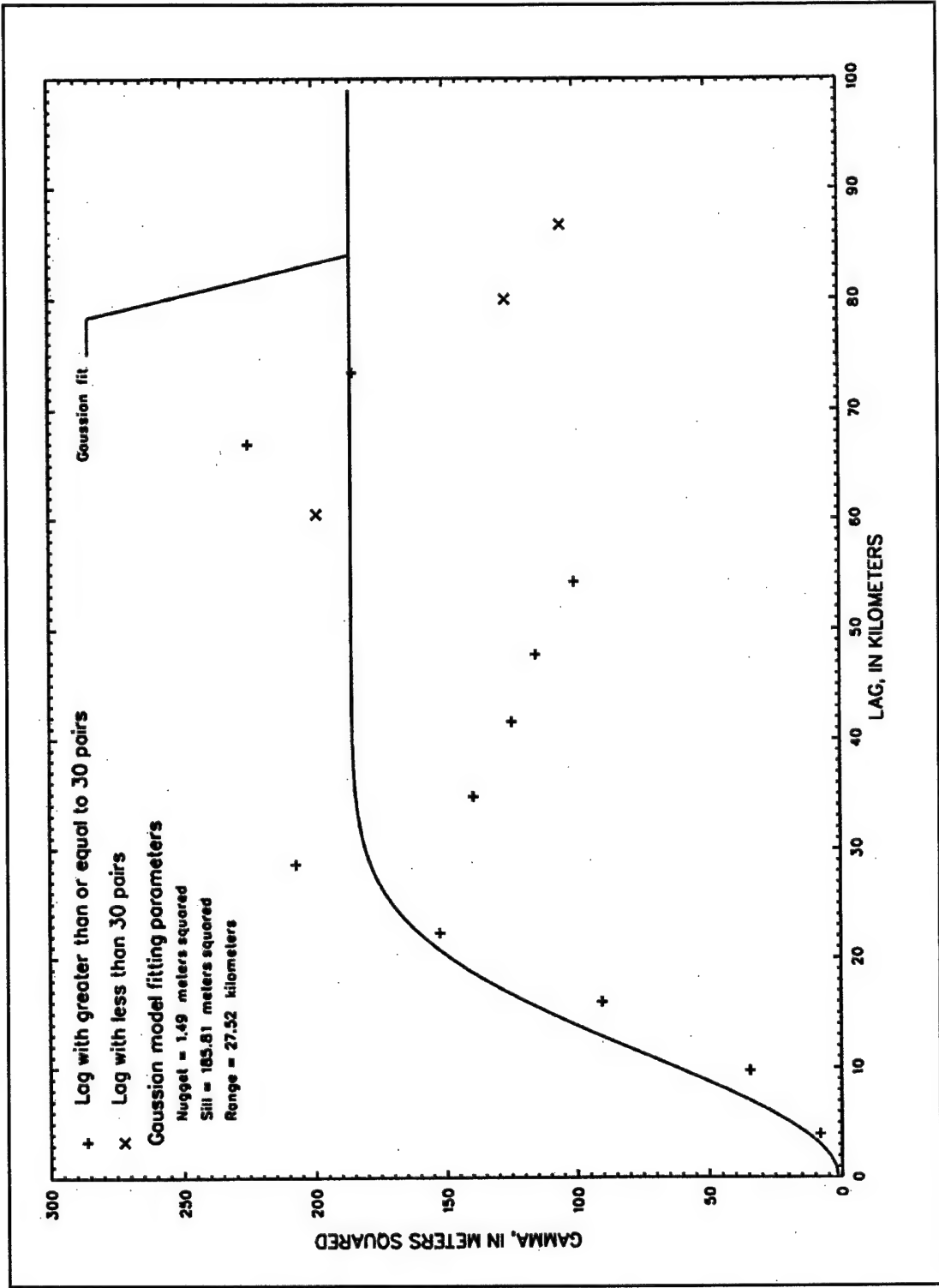


Figure 4-8. Sample variogram points and theoretical Gaussian fit for Iterated Saratoga residuals

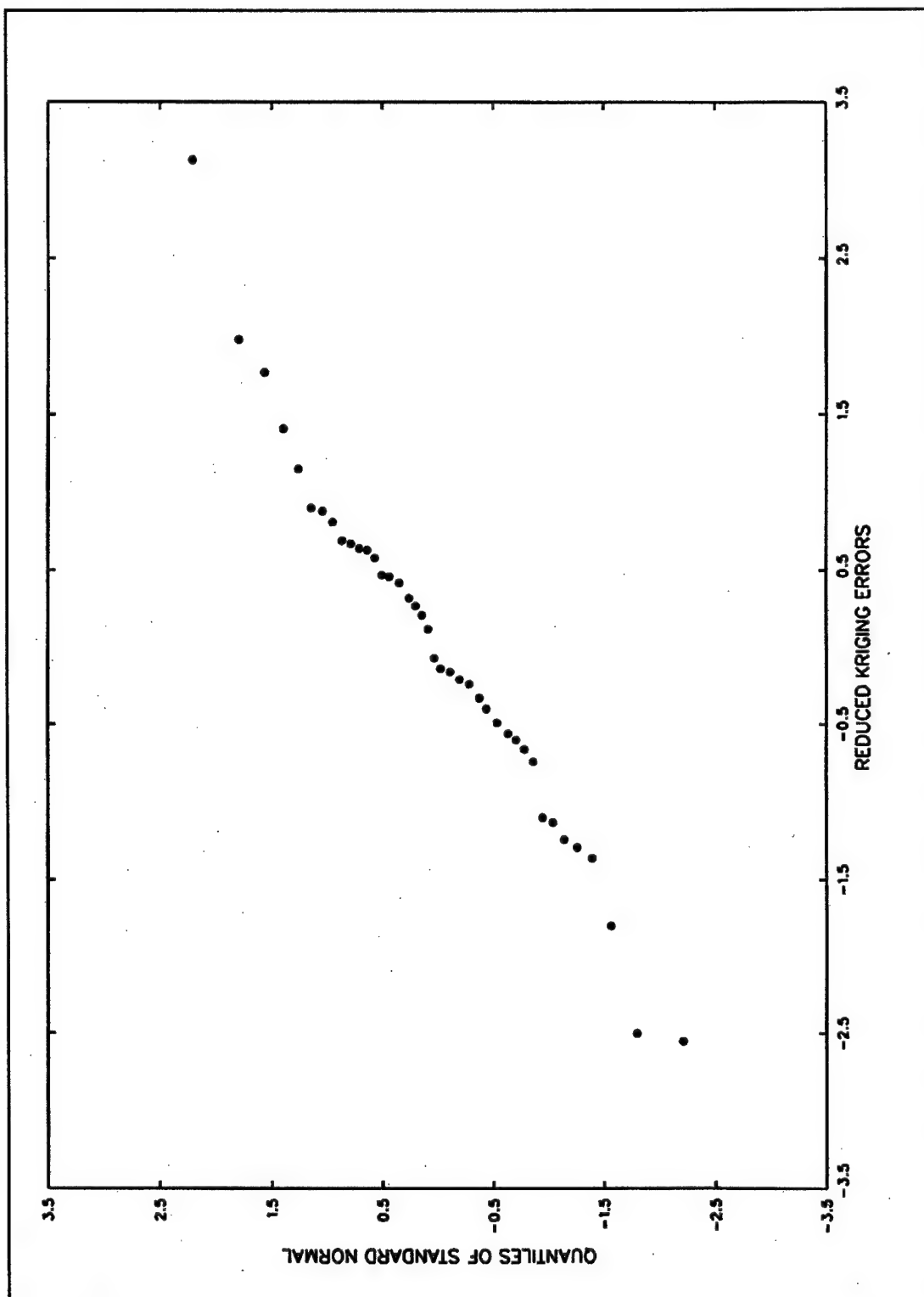


Figure 4-9. Cross-validation probability plot for Saratoga data

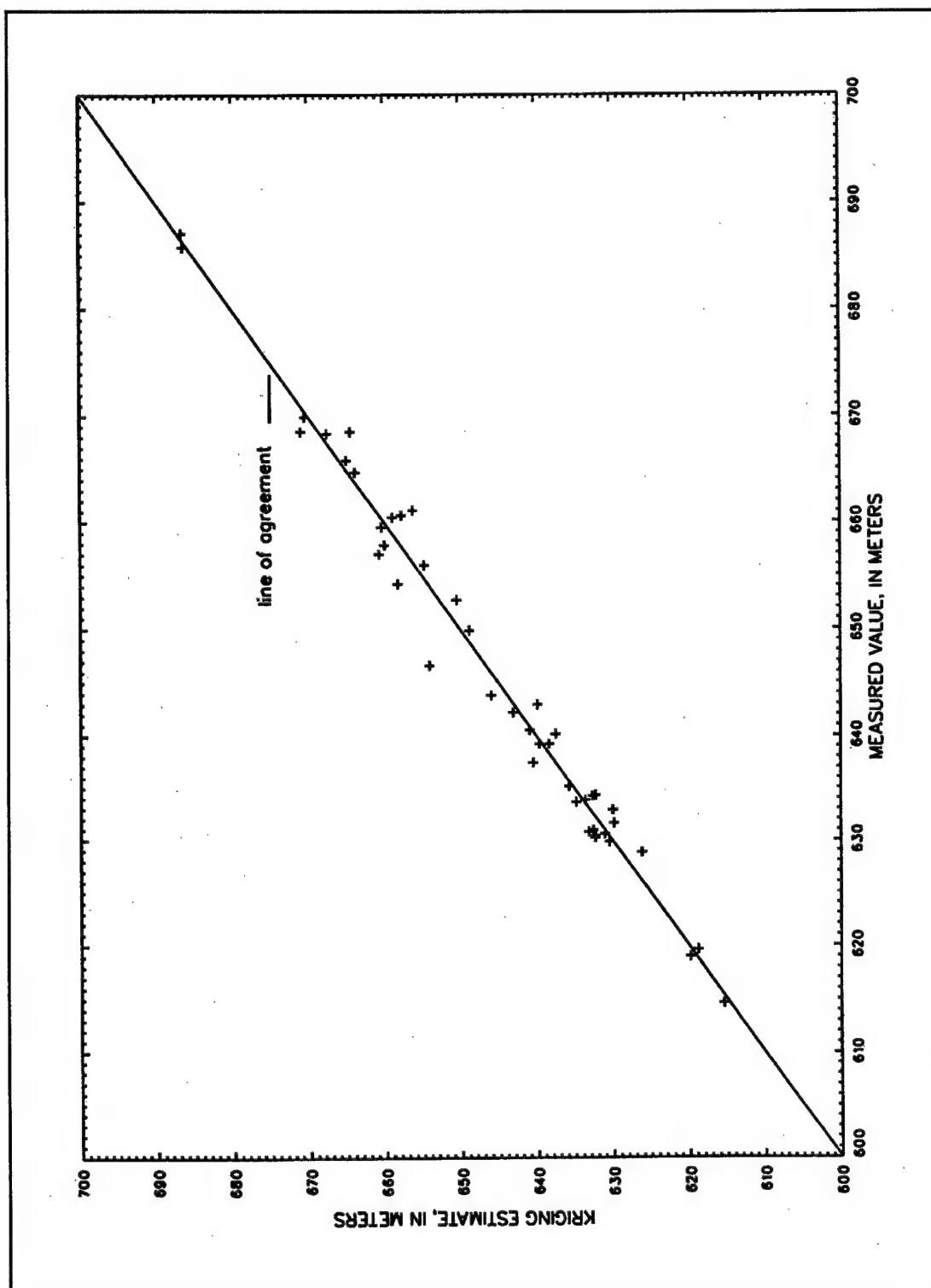


Figure 4-10. Scatter plot of measured versus kriging estimates from cross-validation of Saratoga data

Chapter 5

Practical Aspects of Geostatistics in Hazardous, Toxic, and Radioactive Waste Site Investigations

5-1. General

a. In this chapter, several example applications are described. The applications have been developed using hydrologic, geologic, and contaminant data from established and well-studied hazardous waste sites. The real nature of the data permits discussion of some problems that can occur during HTRW site investigations that stem not only from natural field conditions, but also from typical problems that are associated with the types of data involved. In addition, the real nature of the example data provides an opportunity for comparison between kriging estimates and the real data. In accordance with the purpose and scope of this ETL, these comparisons will be brief and general. This ETL does not provide the comprehensive analysis of data that is addressed by other more elaborate studies.

b. The principal intent of the examples is to provide systematic descriptions for a few of the large number of possible types of applications that investigators may use during HTRW site investigations. The examples are not intended to provide guidance for comprehensive analysis of the included data. This ETL will, however, present some fundamental problems that can occur in geostatistical applications and, in some examples, indicate some possible alternatives.

c. With each example, a purpose will be established and a general environmental setting will be given. Most aspects of variogram construction and calibration will be briefly described and illustrated graphically and in tabular form. A comprehensive treatment of variogram construction has been presented in Chapter 4.

d. GEO-EAS software has been used whenever the example data did not need universal kriging; for those examples, STATPAC was used. As

indicated in Chapter 3, both of these software packages run on the DOS platform (Table 3-1), which will probably be most convenient to readers. The results of kriging estimates are portrayed by gray-scale maps rather than by contours because of the objective nature of the gray-scale format. North is at the top of all maps presented in this ETL, although this orientation may represent some deviation from the real data.

5-2. Water-Level Examples

a. The following examples are for groundwater levels. The principal purpose of the examples is to expose the reader to a kriging exercise using groundwater levels and to indicate how, in a simple manner, kriging standard deviations may be useful to investigators interested in evaluating monitoring networks. The data come from a water-table setting in unconsolidated sediments where the local relief for the land surface is about 30 m. The data involved in this example are considered virtually free of actual measurement error.

b. The location of measured water levels is shown in Figure 5-1a and the basic univariate statistics for this data set are listed in Table 4-1; modifications to the measured data, in the form of addition and removal of measured values, are shown in Figures 5-1b and 5-1c. The techniques described in Chapter 4 were used to guide the following steps for variogram construction:

(1) A raw variogram analysis, along with basic hydrologic knowledge of water-level behavior, indicated that universal kriging would be needed for this analysis.

(2) To obtain a stable variogram of residuals, an iterative, generalized least-squares operation was initially used to remove prominent linear drift of the form $a + bu + cv$, observed in the measured water levels.

(3) After drift was removed, residuals were determined to be stationary and universal kriging with a linear drift was appropriate.

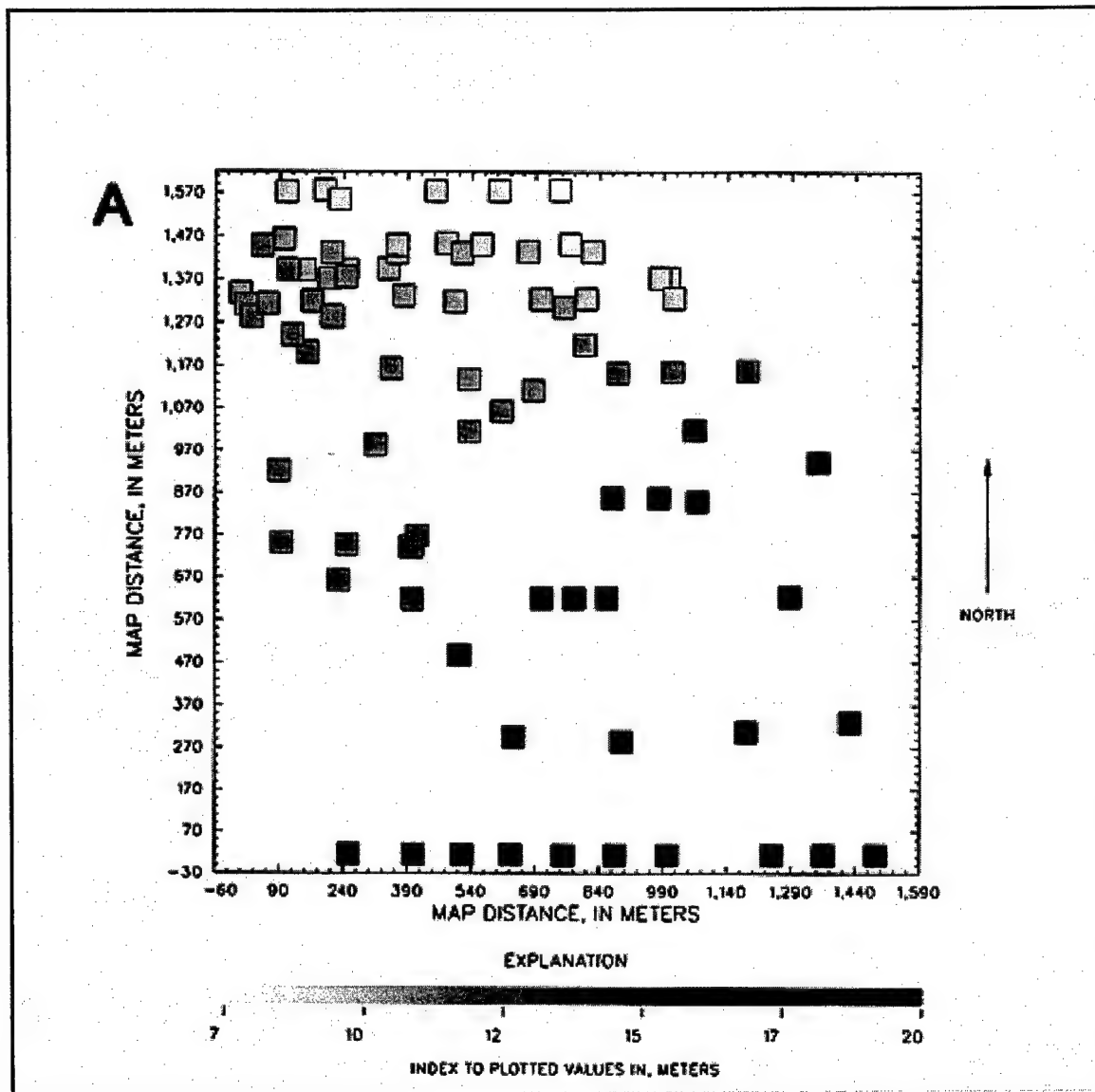


Figure 5-1. Measured data for water-level examples--A, original data; B, original data without dropped sites; C, original data with added sites (added sites indicated with +) (Sheet 1 of 3)

A Gaussian model was used to fit the stabilized variogram of residuals (Figure 5-2a), which has a nugget of 0.093 m^2 , a sill of 2.69 m^2 and a range of 1,219 m (Table 5-1).

(5) Cross-validation was performed, and the results are shown in Figures 5-2b and 5-2c, and

listed in Table 5-1. Cross-validation statistics conform to the criteria discussed in Chapter 4.

c. Linear drift is commonly observed in groundwater elevation data where there are no major anthropogenic activities, such as large groundwater withdrawals. With these

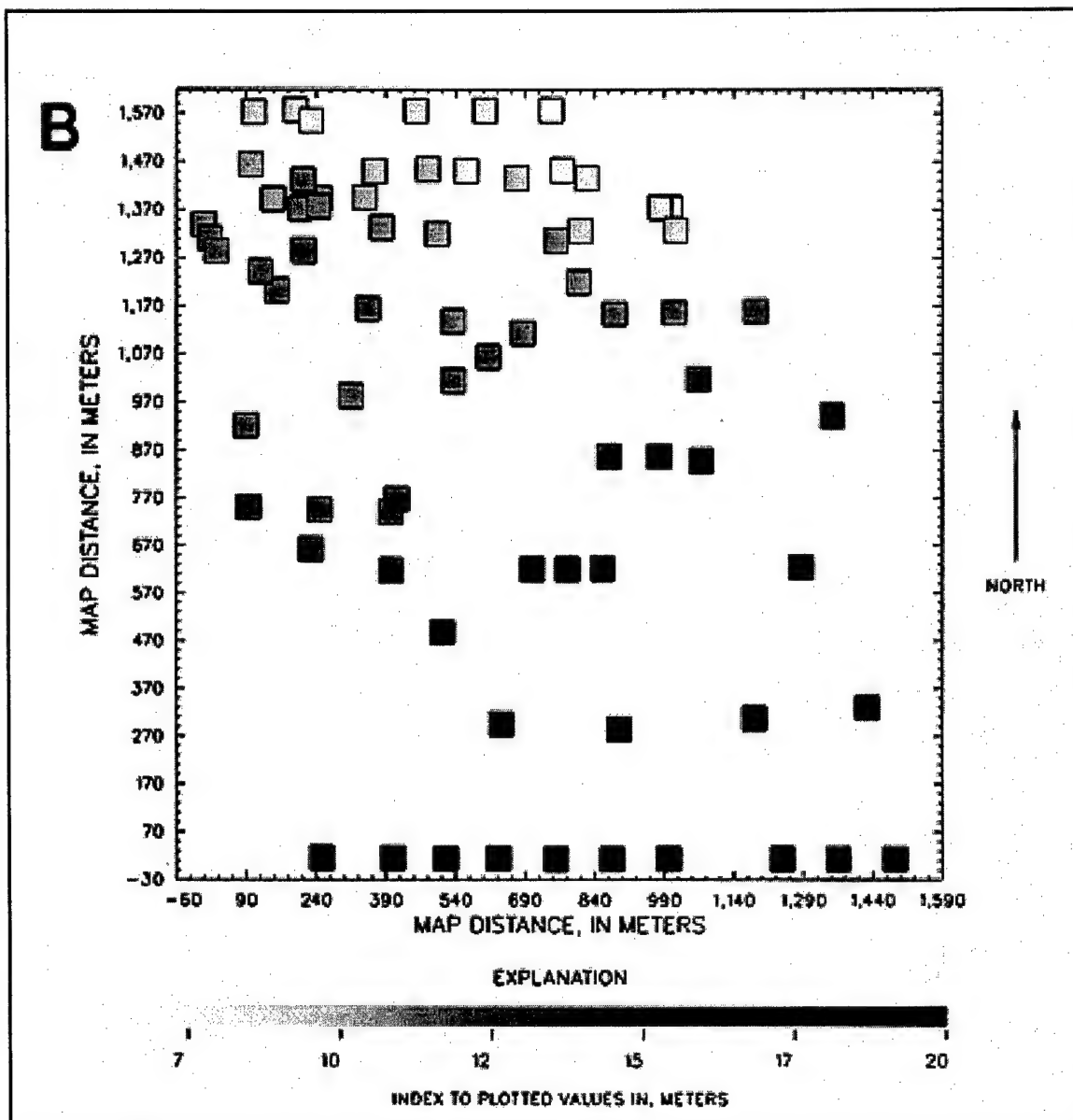


Figure 5-1. (Sheet 2 of 3)

circumstances there is usually a fairly uniform and general groundwater movement that is generally expressed in terms of direction. This uniform and general nature introduces a nonstationary element to the data that, in geostatistics, is referred to as drift. As indicated in Chapter 4, the presence of drift is indicated by a parabolic variogram shape. In this example, the initial variogram in the raw variogram analysis had a characteristic parabolic

shape and a linear drift was identified. Once the drift was identified and characterized, universal kriging procedures were used.

d. A Gaussian model is usually appropriate for variograms with highly continuous variables such as groundwater-elevation data, and it is particularly appropriate in this example. The variogram (Figure 5-2a) at small lags beyond the nugget

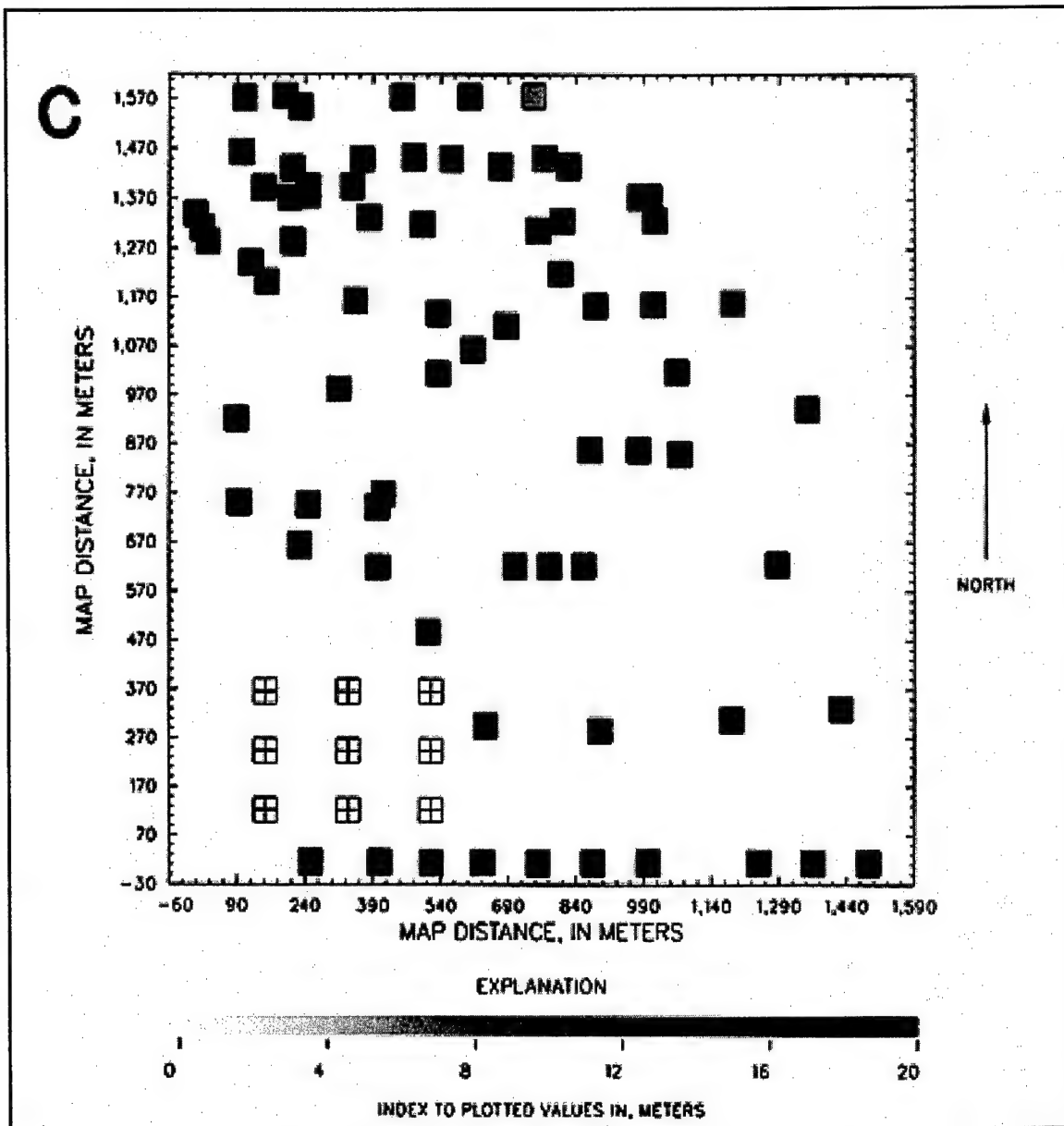


Figure 5-1. (Sheet 3 of 3)

has an upward concavity that cannot be fit with a linear, spherical, or exponential model. The observed shape was interpreted as a function of continuous small-scale variability. The Gaussian model fits the bowl shape of the small lag data (and other data to a lag of about 610 m) well, but it is not flexible enough to closely fit the points much beyond 610 m, indicating that kriging esti-

mates should be computed using neighborhoods with a search radius less than 610 m. In Chapter 4, the initial part of the variogram was described as having the most effect on subsequent kriging estimates.

e. The established variogram then was used with the measured data to produce universal

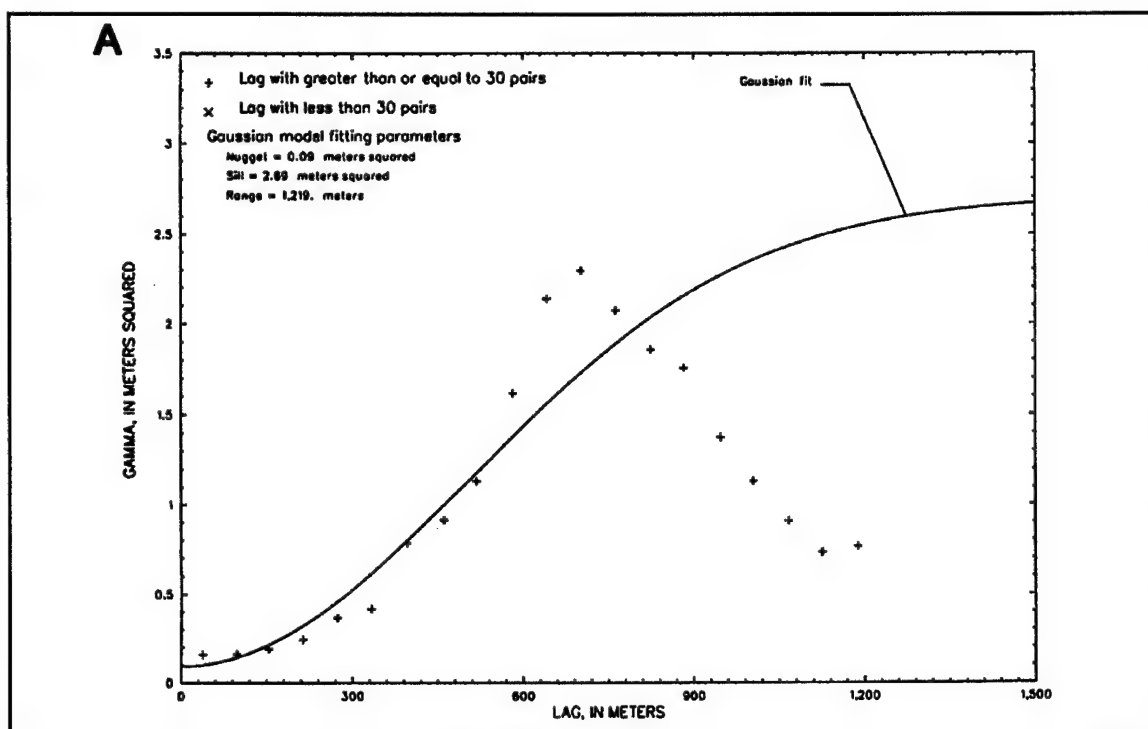


Figure 5-2. Variogram and variogram cross-validation plots for residuals in water-level example--A, theoretical variogram; B, cross-validation scatterplot; C, cross-validation probability plot (Sheet 1 of 3)

kriging estimates for all points in a 26-by-26 grid with a grid size of about 61-by-61 m. A gray-scale map of the kriging water levels is shown in Figure 5-3a and basic univariate kriging estimate statistics are listed in Table 5-2a (water level A). The kriging results are a good representation of the results from other more elaborate studies.

f. Kriging standard deviations for the kriging estimates are shown in Figure 5-3b. The magnitude of kriging standard deviations can provide investigators with a direct indication of where the uncertainty associated with kriging estimates is relatively high or low. The areas of greatest uncertainty for the kriged water levels are in the upper right and lower left corners of the map, where standard deviations are as high as about 1.4 and 0.8. Not surprisingly, these areas are where the density of the measured data is relatively low. Throughout much of the remainder (about 70 percent) of the map, the kriging standard deviation is almost constant at about 0.35.

g. To use the kriging standard-deviation values in a more quantitative manner, the investigator needs to establish some assurance that the measured data and the reduced kriging errors are approximately normally distributed and also that the assumption of stationary residuals after drift removal is correct. If the investigator is confident about these assumptions, then the basic statistical principles involving confidence intervals can be applied. In this example, the standard deviation of about 0.35 throughout most of the map indicates that there is a 95-percent chance that the true value at a location where there is a kriging estimate will be within about 0.70 (twice the kriging standard deviation) of the kriging estimate.

h. As an example of evaluating network density and the accuracy of kriging estimates, two new maps were developed. To make the first map, a decrease in network density was effected by removing nine measured locations from the north-west part of the area (Figure 5-1b) where sampling

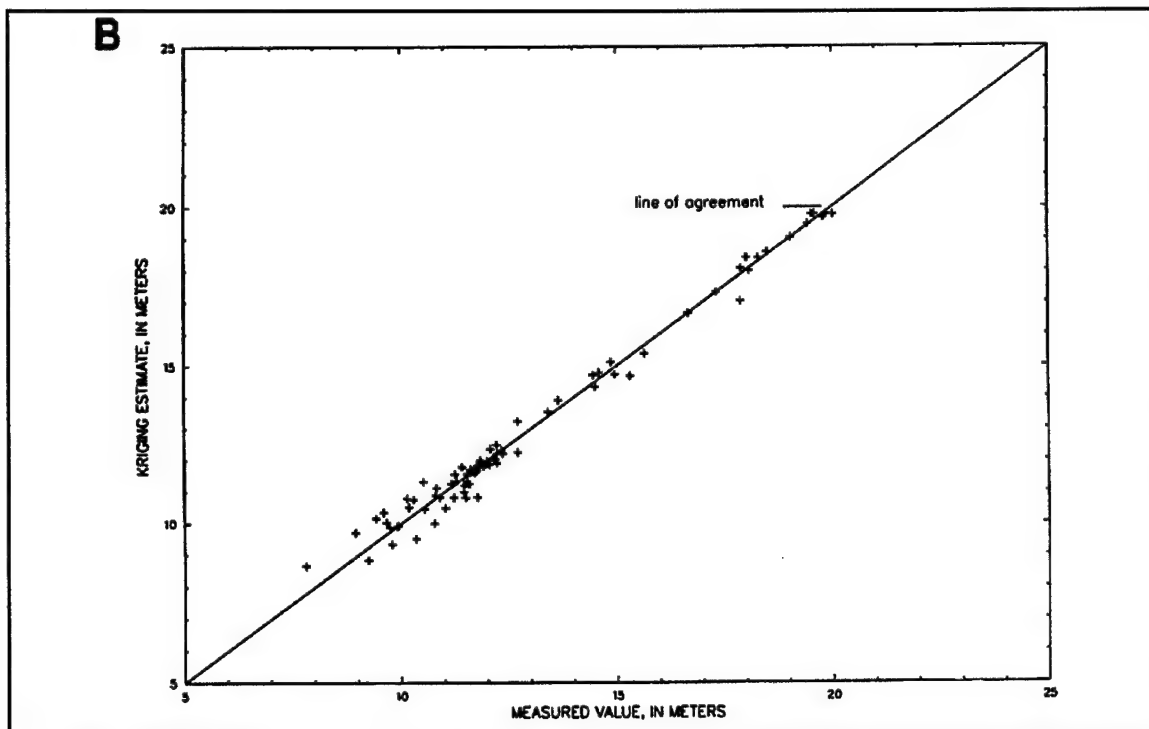


Figure 5-2. (Sheet 2 of 3)

density was high and kriging standard deviations were low. Kriging estimates were produced for the same grid and the basic univariate kriging estimate statistics are listed in Table 5-2 (water level B). The map shown in Figure 5-3c indicates that the ratio of the original kriging standard deviations and the kriging standard deviations with the nine measured locations removed is always very close to 1.00, which indicates that there is very little difference between the two sets of kriging standard deviations and that water levels are oversampled in the area where the nine measured locations were removed.

i. To produce the second map (Figure 5-1c) nine locations were added in the southwest corner where the sampling density was relatively low and the kriging standard deviation was relatively high. In section 2-4, Equation 2-47 indicates that the universal kriging variance depends on the variogram, the type of trend, and measurement locations; in this respect the kriging standard deviation does not depend on the values at measurement

locations. Consequently, values of zero were used for the nine new measurement locations and only the resultant map of kriging standard deviations (Figure 5-3d) is of interest. The map shows that the kriging standard deviations in the lower left corner, which formerly had values of about 0.8, have been decreased by a factor of approximately 0.25, which indicates that the kriging estimates, based on the geometry of the network, are more reliable.

5-3. Bedrock-Elevation Examples

a. The following examples are for bedrock elevations. The principal purposes of the examples are to familiarize the reader with a kriging exercise using bedrock elevations and to describe block kriging. The data come from an area where bedrock consists of a series of intercalated terrestrial deposits that have been weathered somewhat and then covered with alluvium. The opportunity for measurement error in these types of data is

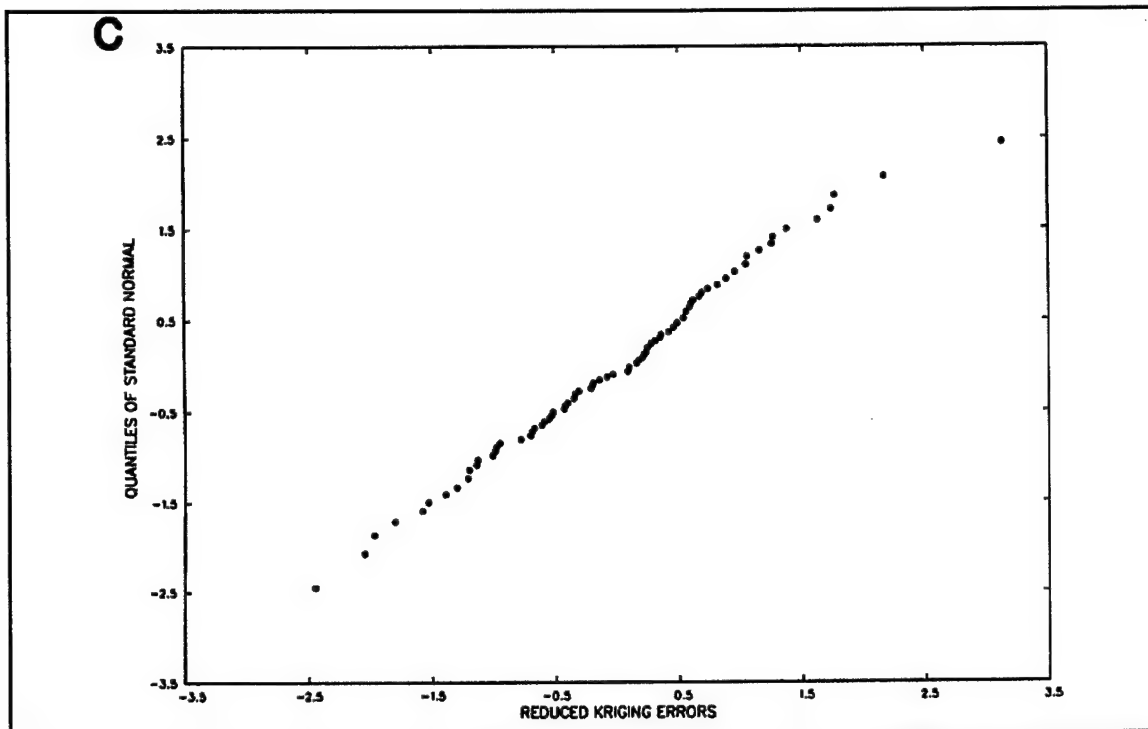


Figure 5-2. (Sheet 3 of 3)

inevitable because the determination of just where bedrock begins is complicated and subjective.

b. The set of measured locations, set A, is shown in Figure 5-4a and the basic univariate statistics are listed in Table 4-1 (bedrock A); modifications to the measured data, in the form of removal of sites is shown in Figure 5-4b. The techniques described in section 4-1 were used to guide the following steps for variogram construction:

(1) The raw variogram indicated a stationary spatial mean. The data were assumed to be suitable for ordinary kriging.

(2) An isotropic Gaussian model was used to fit the variogram which had a nugget of 0.650 m^2 , a sill of 12.54 m^2 , and a range of 914 m (Table 5-1, bedrock A).

(3) Cross-validation was performed, and the results, (Table 5-1, bedrock A), were not acceptable.

c. The cross-validation exercise produced a reduced-root-mean-squared error of 2.146 [Table 5-1 (bedrock A)] which indicates, as described in Chapter 4, that the kriging variance is underestimated to an unsatisfactory degree. Further attempts to fit the Gaussian model to the sample variogram points produced better cross-validation statistics; however, the Gaussian curve began to depart substantially from the sample variogram points at the lower lag sample points. As a result, the distribution of the residuals was explored, and the eastern, and especially north-eastern, parts were determined to contain problematic data values that rendered the distribution nonhomogeneous. The nonhomogeneous nature is related to an incised channel present on the

Table 5-1
Variogram Characteristics and Cross-Validation Statistics¹

Example Identifier	Variogram Characteristics					Cross-Validation Statistics			
	Transformation	Direction/ tolerance	Model	Nugget (Base units squared)	Sill (Base units squared)	Range (Base units)	Average Kriging Error (Base units)	Kriging Root- Mean-Squared Error (Base Units)	Reduced Root- Mean-Squared Error (dimensionless)
Water levels	Drift	0/NA	Gaussian	1.00	29.00	4,000.00	-0.002	1.20	1.083
Bedrock A	None	0/NA	Gaussian	7.00	135.00	3,000.00	0.146	8.30	2.146
Bedrock B	None	0/NA	Gaussian	8.00	90.00	2,400.00	-0.034	4.39	1.192
Water Quality A	Natural log	150/45	Exponential	1.00	3.20	4,250.00	0.105	1.54	0.938
Water Quality A	Natural log	240/45	Exponential	1.00	3.20	750.00	0.105	1.54	0.938
Water Quality B	Indicator	150/45	Spherical	0.05	0.25	2,000.00	NA	NA	NA
Water Quality B	Indicator	240/45	Spherical	0.05	0.25	700.00	NA	NA	NA

¹NA= not applicable; base unit for water levels and bedrock is feet, base unit for water quality A is log concentration, concentration in micrograms per liter, base unit for water quality B is indicator units.

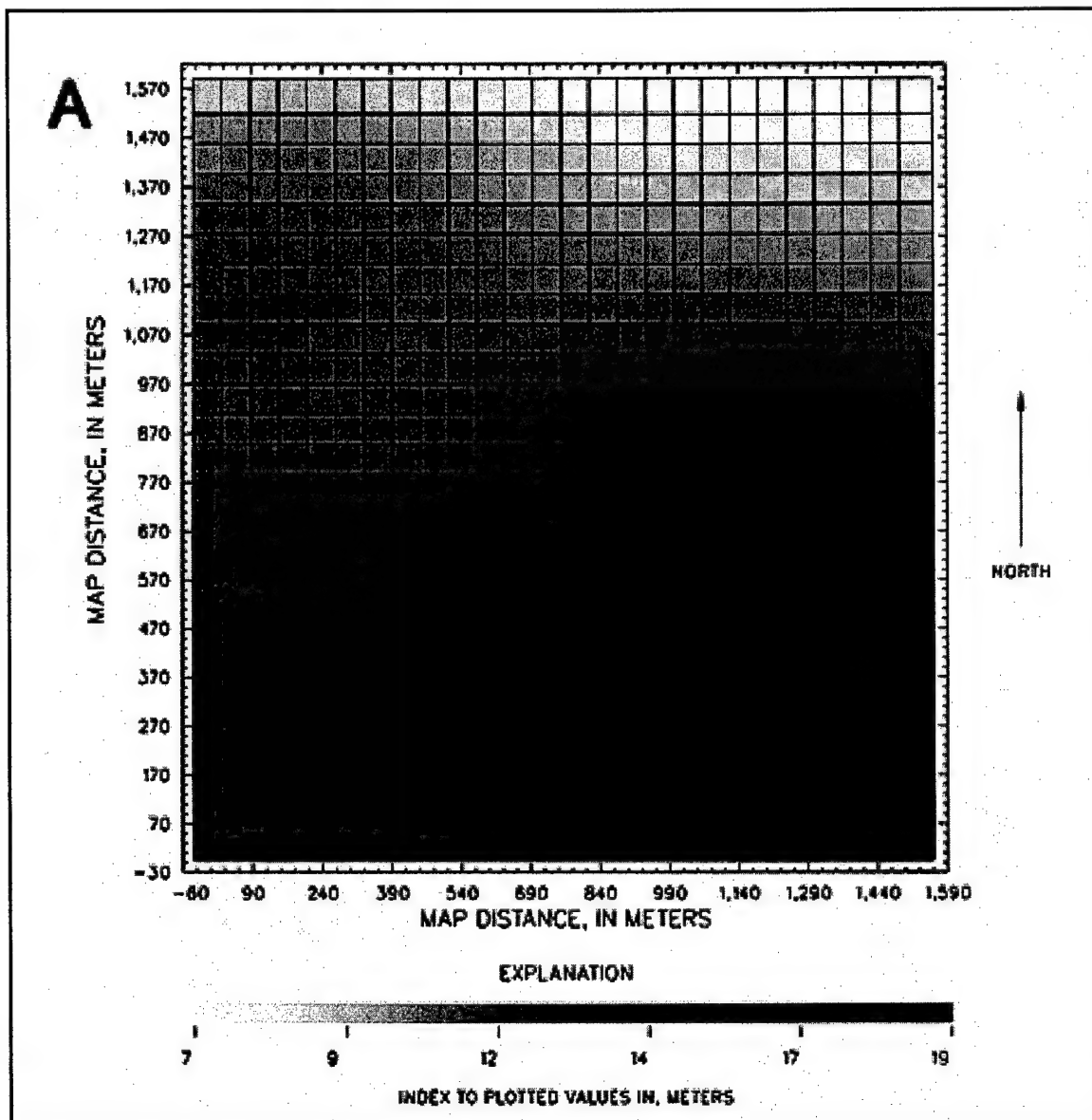


Figure 5-3. Kriging results for water-level examples--A, kriging estimates for original data; B, kriging standard deviations for original data; C, ratio (original data to original with dropped sites) of kriging standard deviations; D, kriging standard deviations for original data with added sites (Sheet 1 of 4)

bedrock surface. At this juncture, the measured data were restricted to exclude the outlying measurements. Before this decision was made, two alternative methods for dealing with the outlying values were considered and deemed beyond the scope of this effort. However, a brief discussion of the situation is appropriate.

d. The first alternative considered was to fit a contrived and nongradual surface to the measured data and remove the outlier effect. A splined surface could be capable of producing the desired result. The decision whether or not to pursue such a remedy becomes somewhat philosophical. In a relatively simple example, as in this bedrock

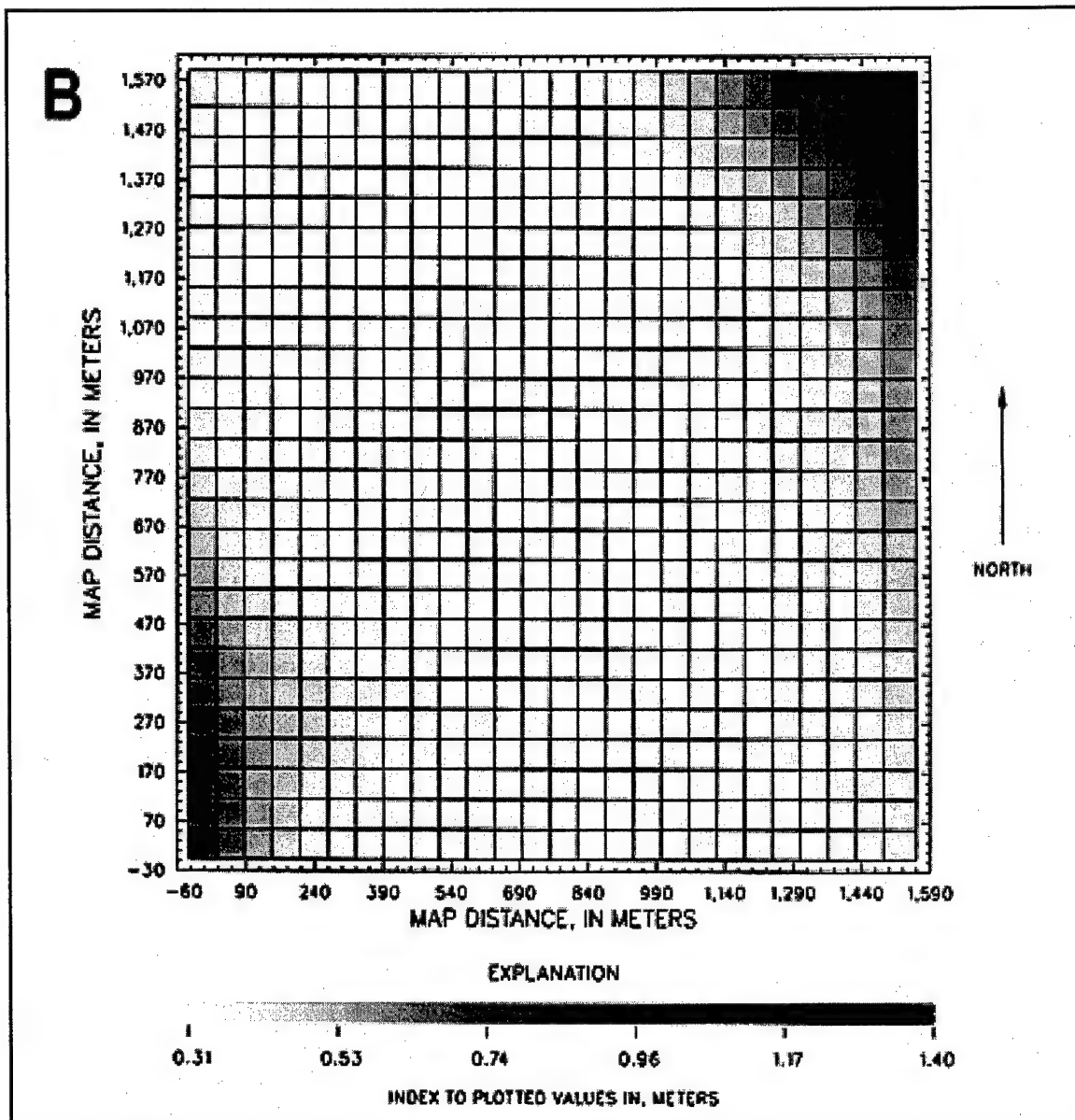


Figure 5-3. (Sheet 2 of 4)

example, such a remedy may be entirely appropriate; however, some investigators may support the idea that the situation is actually dealing with two unique and homogeneous domains. Therefore, the second alternative considered, distributing the kriging process so that each homogeneous domain is addressed independently, becomes more attractive. In more complicated applications where a large number of domains are present, a distributed

approach may be necessary to avoid an undue amount of compromise.

e. The restriction of measured data, set B, is shown in Figure 5-4b and the basic univariate statistics are listed in Table 4-1 (bedrock B). The restriction exercise resulted in removing 17 measured locations and in the truncation of the north-eastern part of the area so that the area became

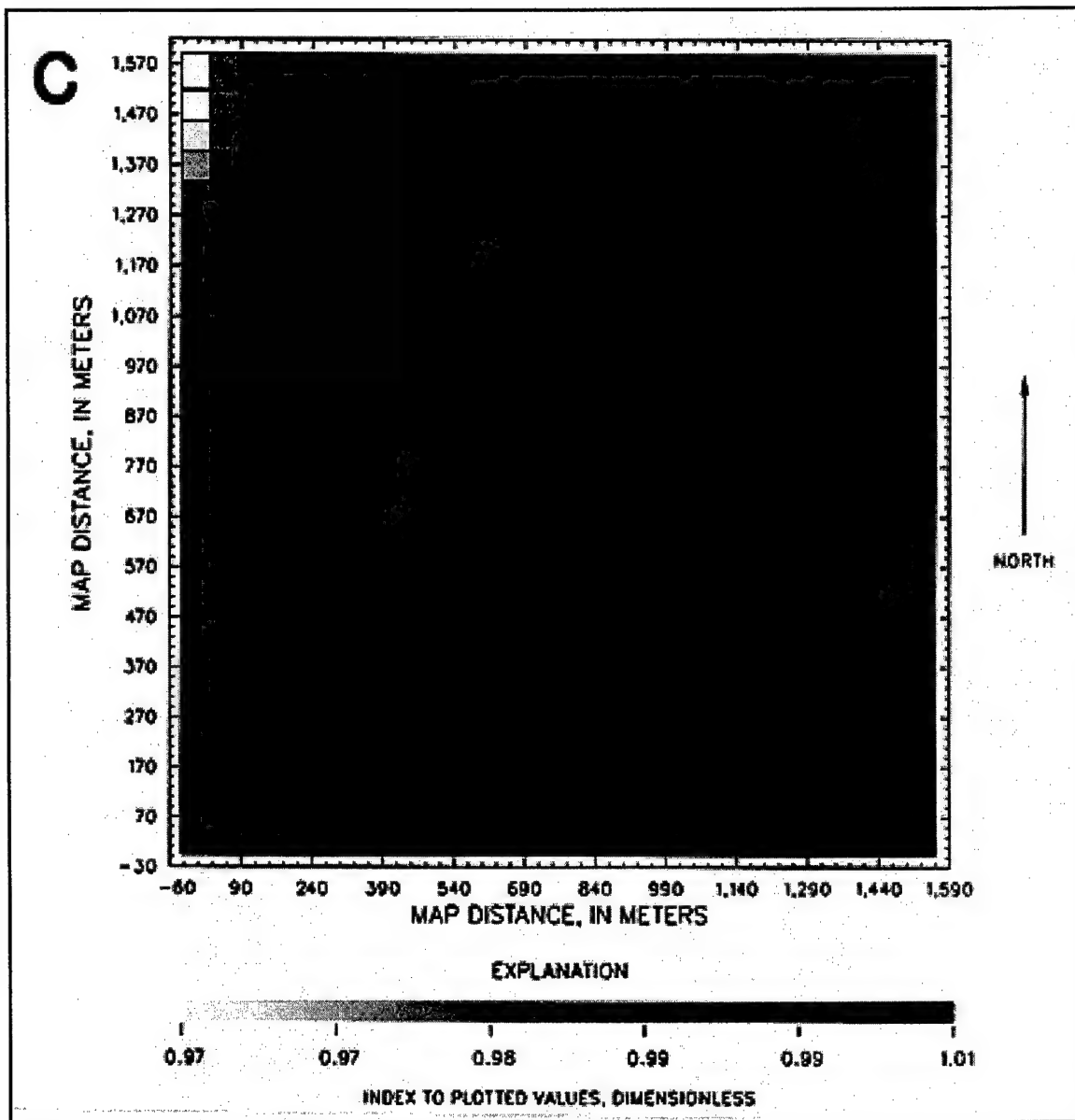


Figure 5-3. (Sheet 3 of 4)

polygonal rather than rectangular. Again, the techniques described in Chapter 4 were used to guide the following steps for variogram construction:

(1) A Gaussian model was used to fit the variogram which had a nugget of 0.650 m^2 , a sill of 8.36 m^2 , and a range of 732 m. The variogram indicated a stationary spatial mean.

(2) Initial cross-validation was performed, and the nugget was changed from 0.650 m^2 to 0.743 m^2 to improve cross-validation statistics. The final variogram is shown in Figure 5-5a and characteristics are listed in Table 5-1.

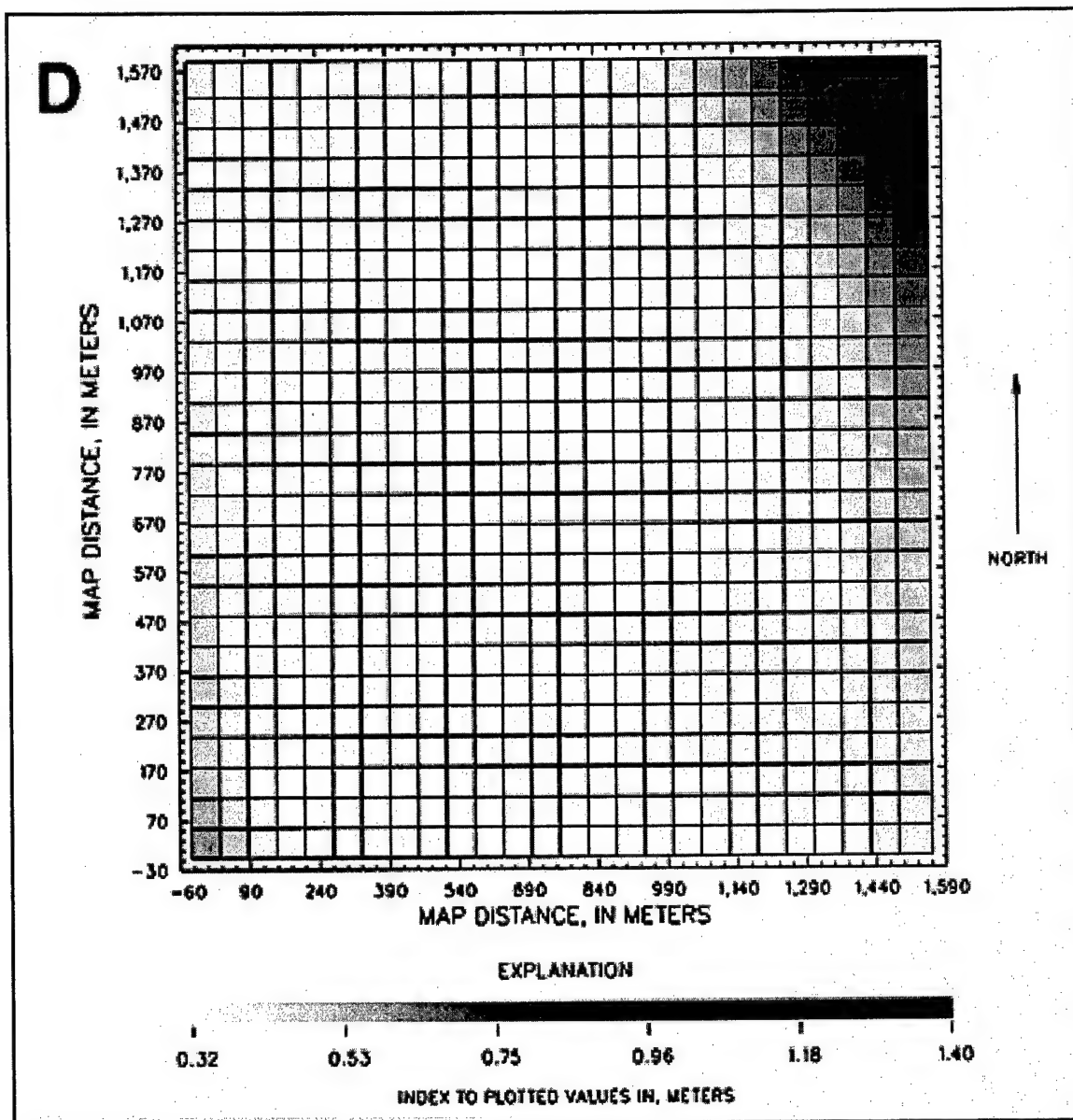


Figure 5-3. (Sheet 4 of 4)

(3) Final cross-validation was performed, and the results, shown on Figures 5-5b and 5-5c and listed in Table 5-1 (bedrock B), were acceptable.

f. The large difference between the sill defined for the initial data set and the sill for the restricted data set (12.54 m^2 and 8.36 m^2) supports the hypothesis that the original data set is actually two different domains. The final variogram then

was used, along with the measured data, to produce ordinary kriging estimates for all points in a 52-by-52 grid with a spacing of about 30-by-30 m, which is truncated along the northeastern border because of the restriction operation. For the kriging procedure, a search radius of about 914 m with a maximum of 16 and minimum of 8 surrounding locations was specified. Gray-scale maps of the kriging estimates and kriging standard

Table 5-2
Univariate Statistics for Gridded Kriging Estimates in Example Applications¹

Example Identifier	Transformation	Minimum (Base units)	Maximum (Base units)	Mean (Base units)	Median base units)	Standard deviation (base units)	Skewness (dimensionless)
Water level A	Drift	24.34	65.00	45.86	44.46	10.15	0.11
Water level B	Drift	24.59	65.00	45.84	44.45	10.14	0.11
Bedrock B	None	26.13	64.88	41.45	39.78	7.71	0.82
Bedrock C	None	26.72	64.39	41.50	39.69	7.63	0.82
Water quality A	Natural log	2.92	7.07	5.17	5.03	0.72	-0.06

¹Base units for water level A and B and bedrock B and C is feet; base unit for water quality A is log concentration, concentration in micrograms per liter.

deviations are shown in Figures 5-6a and 5-6b, respectively, and the univariate kriging estimate statistics are listed in Table 5-2 (bedrock B). The kriging results indicate channel-like features in the bedrock surface, as well as a prominent bedrock high at the south border of the area; the results are a good representation of the results from other more elaborate studies.

g. For an example of block kriging, an investigative goal of establishing block values of bedrock elevation for a finite-difference groundwater model grid having about 120- by 120-m cells was assumed. The same variogram and search criteria were used to estimate block values for a 13-by-13 grid with about 120- by 120-m spacing; a 4-by-4 block was specified. Each kriging value shown in Figure 5-6c is interpreted as an estimate of the average value of bedrock elevation over the about 120- by 120-m block. The standard deviation for the block estimates is less than the standard deviation for the point estimates (Table 5-2). Gray-scale maps of the kriging estimates and the kriging standard deviations are shown in Figures 5-6c and 5-6d, and the univariate kriging estimate statistics are listed in Table 5-2 (bedrock C).

5-4. Water-Quality Examples

a. The following examples are for water-quality information consisting of concentrations

determined for a contaminant. The principal purposes of the examples are to familiarize the reader with a kriging exercise using water-quality information and to illustrate indicator kriging. The examples also will familiarize the reader with data that are strongly anisotropic and need transformation. The data come from a water-table aquifer developed in alluvial sediments where the depth to water was less than about 23 m. Several analytical laboratories were involved in measuring the concentration of the contaminant in the water-quality examples. Each of the analytical laboratories was required to follow rather comprehensive guidelines that specified tests of instrument performance before sample determinations were made, as well as measurement of extraction efficiencies. Because of these performance guidelines, the opportunity for errors due to instrument error was considered to be either known or relatively low. In addition to using performance guidelines, field quality-assurance samples were also collected. These samples can be used to evaluate other types of possible errors, such as cross-contamination and representativeness of the sample. Duplicate samples for the contaminant in the water-quality examples indicate as much as 15 percent variability in reported results. This variability is not entirely unusual and is most likely related to the integrity of the analytical method or the method in which the sample media was aggregated during sample collection.

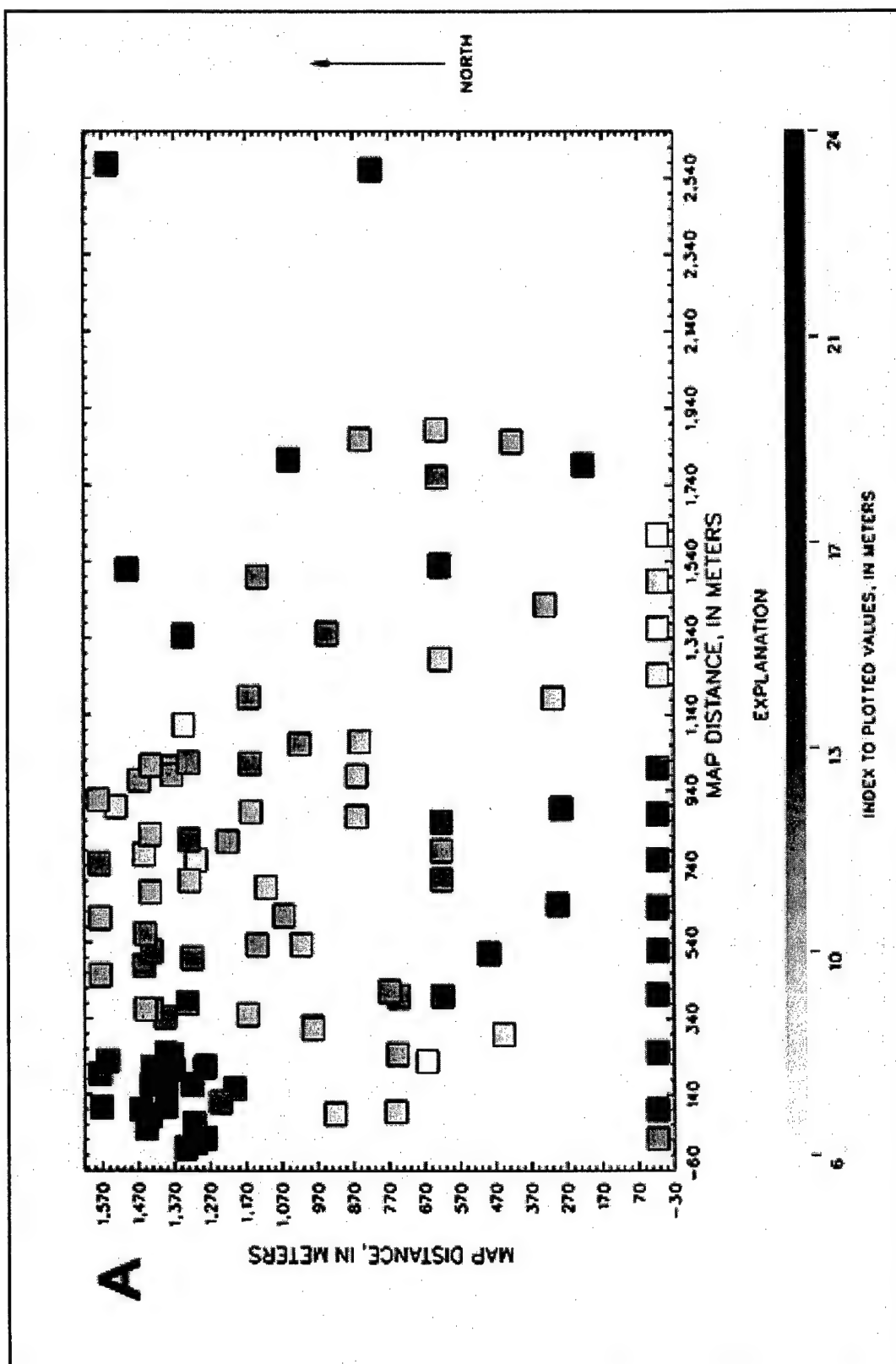


Figure 5-4. Measured data for bedrock-elevation examples--A, original data and B, restricted data (Continued)

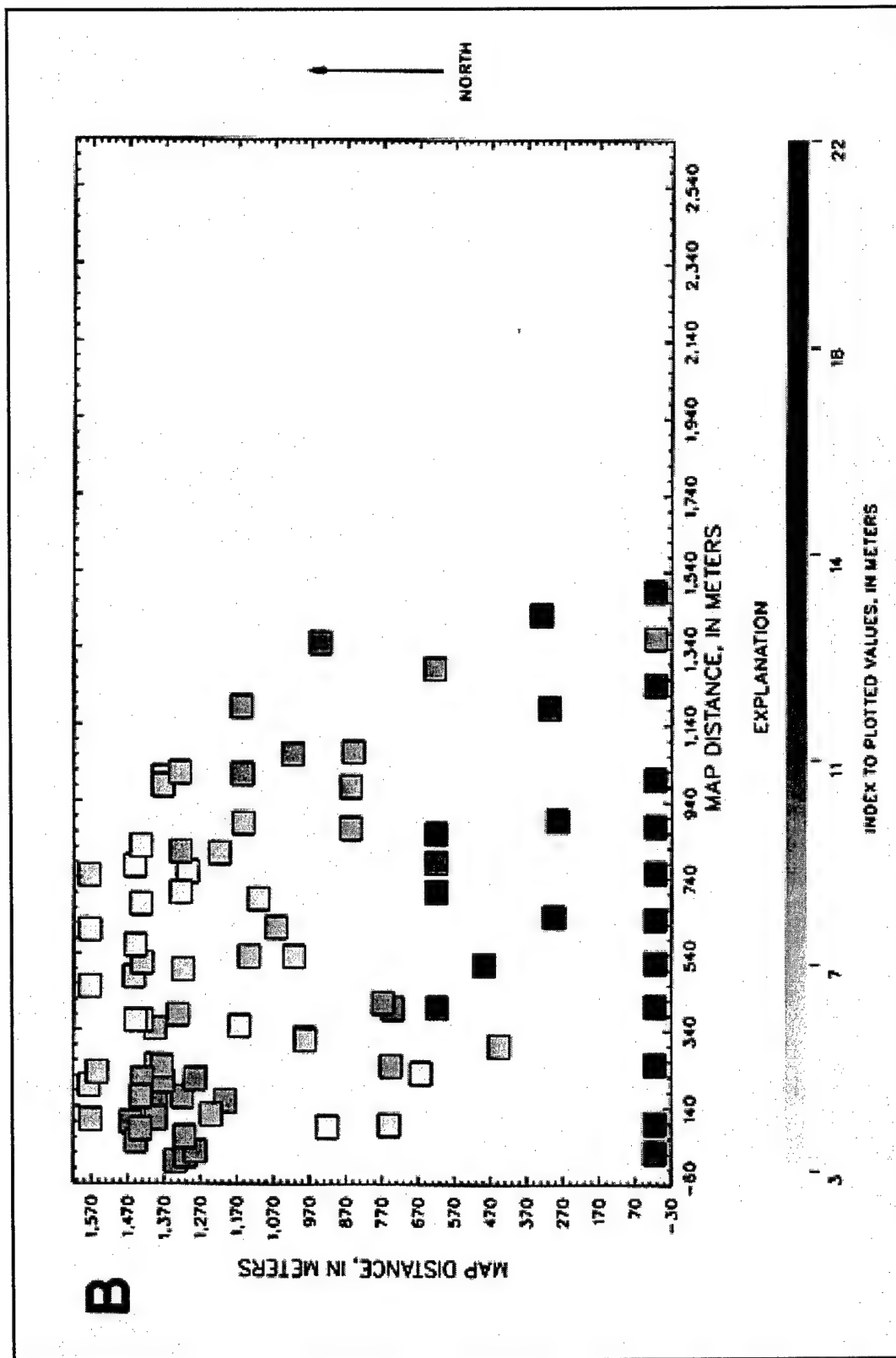


Figure 5-4. (Concluded)

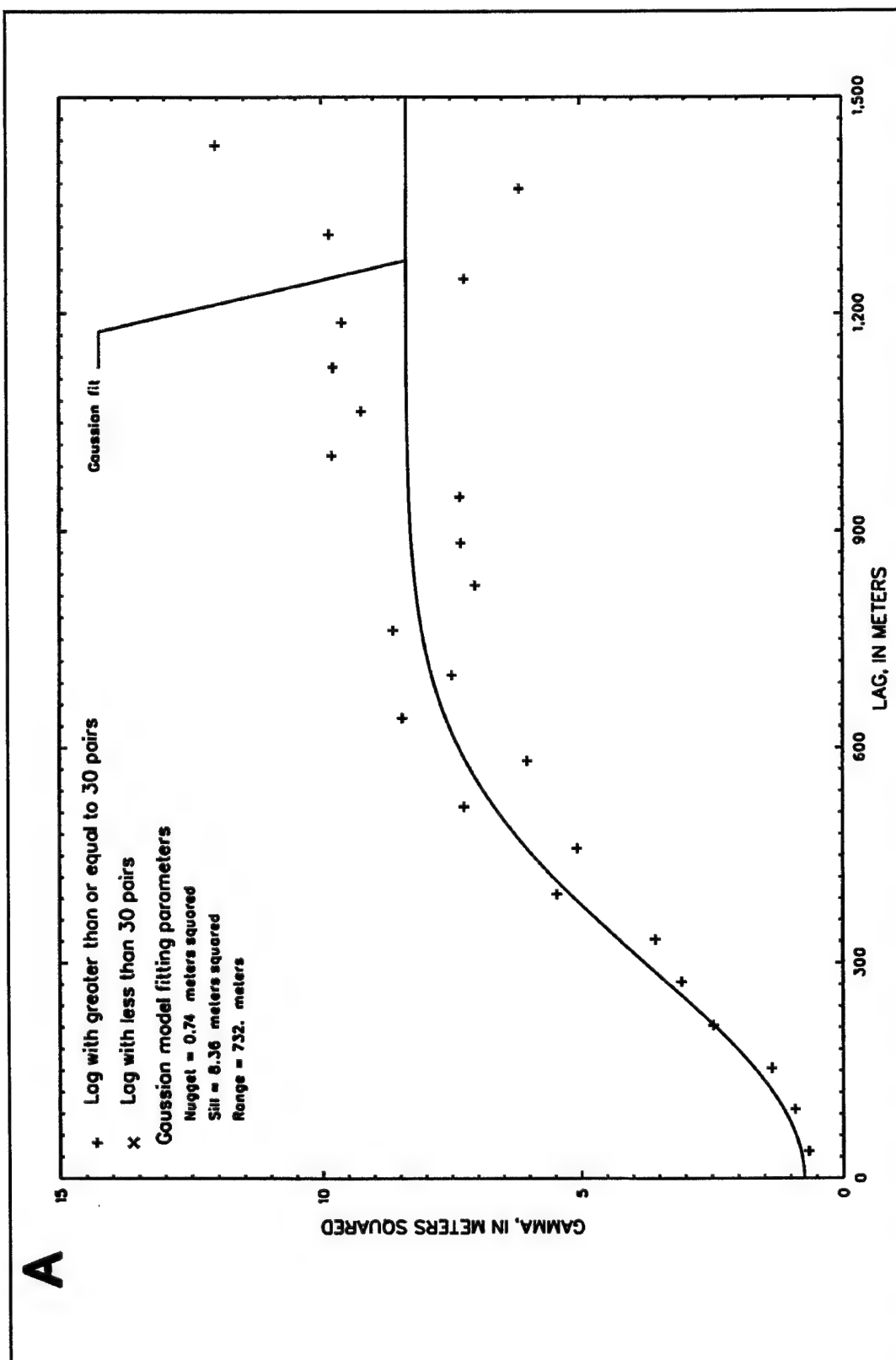


Figure 5-5. Variogram and variogram cross-validation plots for bedrock-elevation example--A, theoretical variogram; B, cross-validation scatterplot; C, cross-validation probability plot (Sheet 1 of 3)

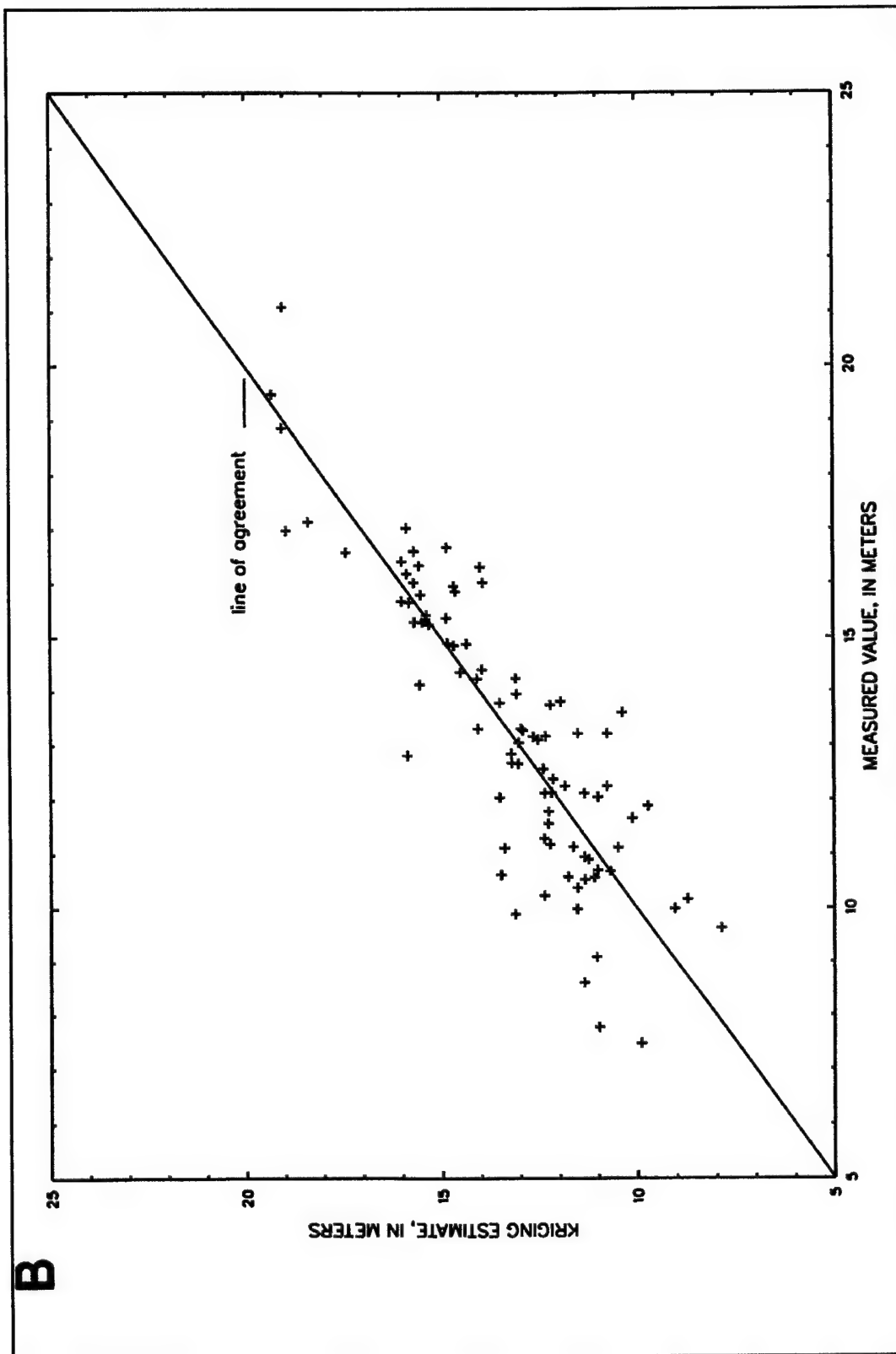


Figure 5-5. (Sheet 2 of 3)

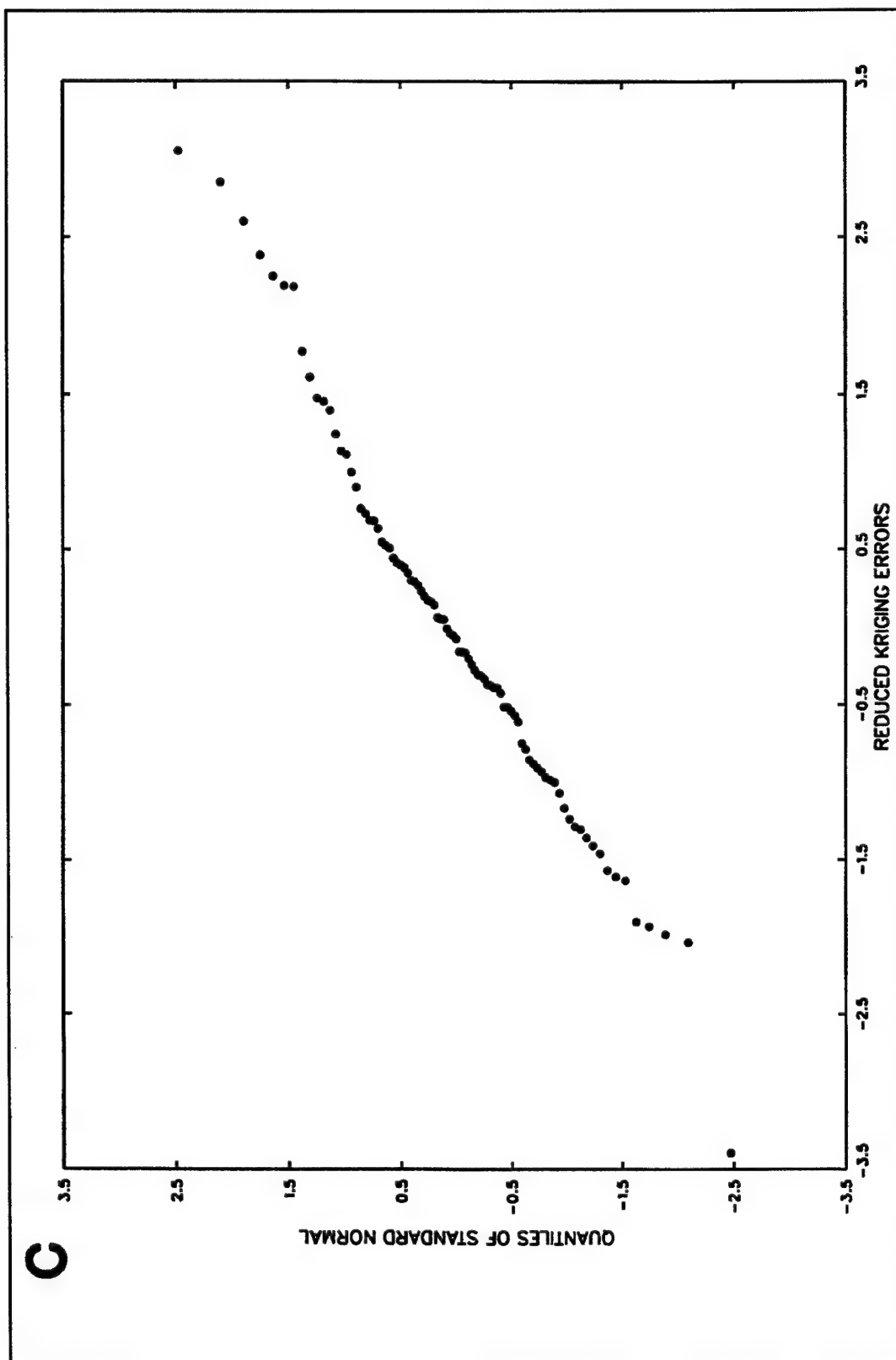


Figure 5-5. (Sheet 3 of 3)

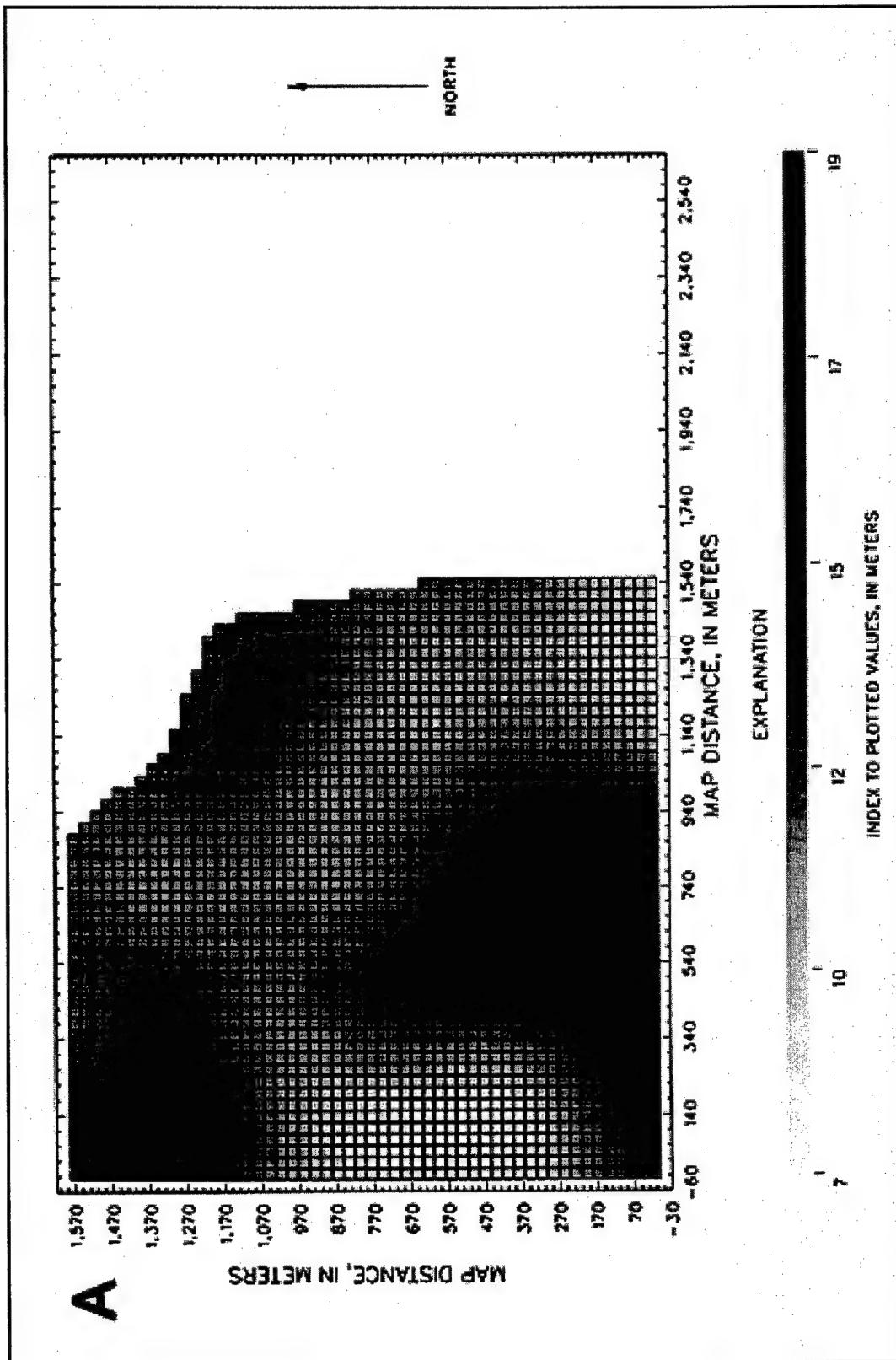


Figure 5-6. Kriging results for bedrock-elevation examples--A, kriging estimates; B, kriging standard deviations; C, block kriging results; D, block kriging standard deviations (Sheet 1 of 4)

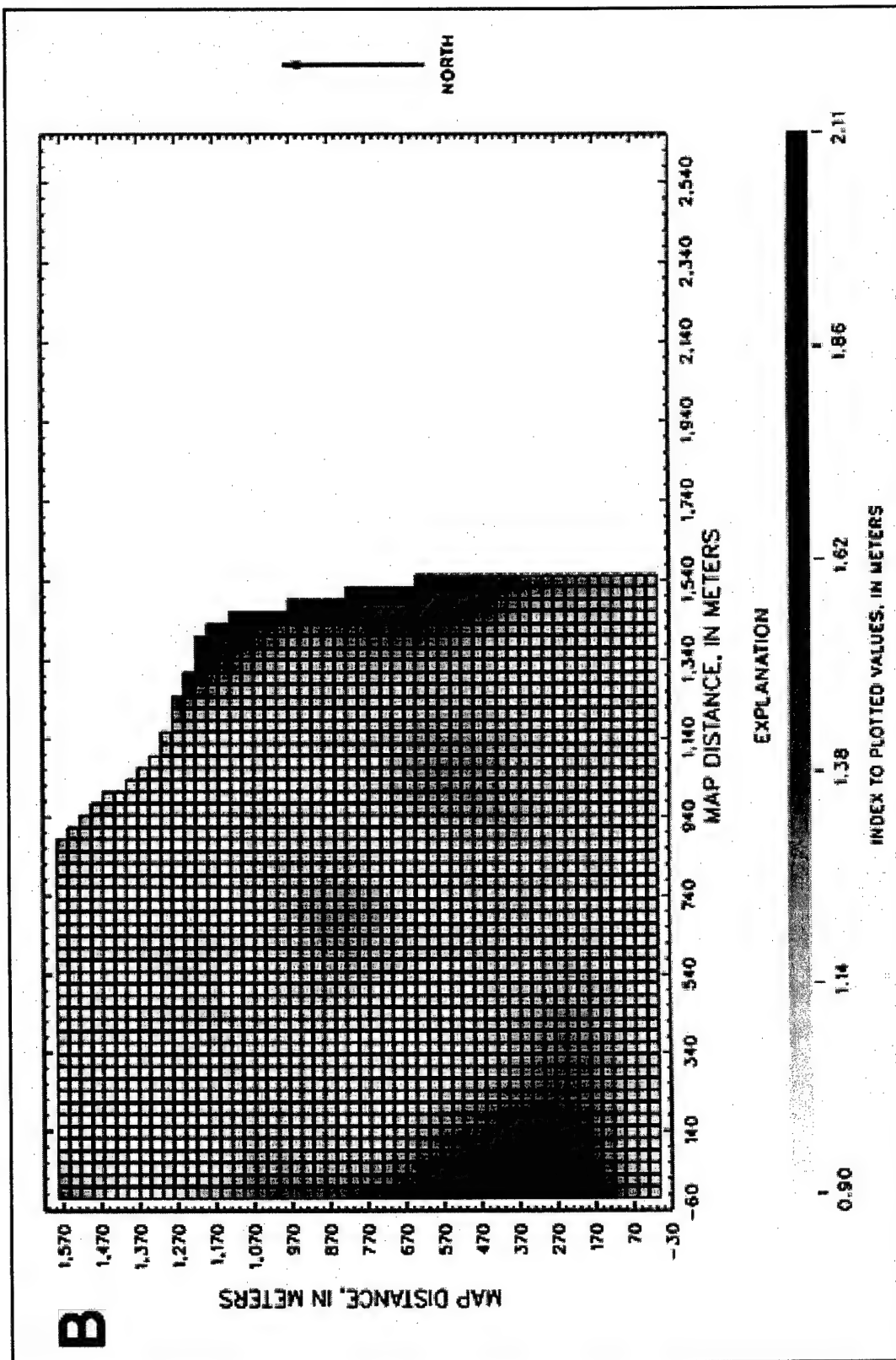


Figure 5-6. (Sheet 2 of 4)



Figure 5-6. (Sheet 3 of 4)

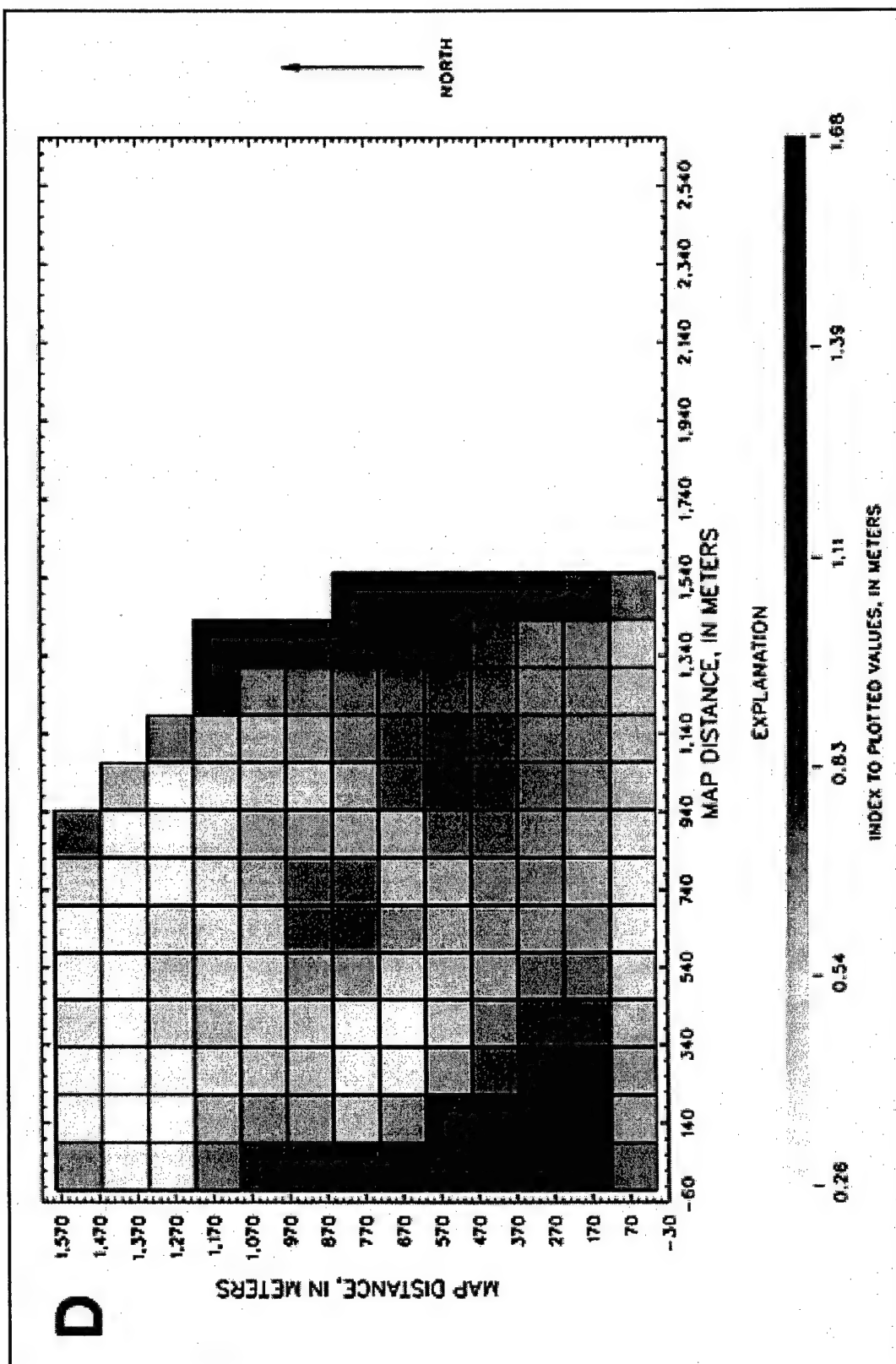


Figure 5-6. (Sheet 4 of 4)

b. Measured locations are shown in Figure 5-7 and the basic univariate statistics are listed in Table 4-1 (water quality A). An initial review of the data indicated three important features.

(1) The data seemed to have strong anisotropy at about 150 counterclockwise degrees to the east-west baseline.

(2) The data required a natural log transformation so the distribution was approximated by a normal distribution.

(3) No trends were indicated during preliminary exploration, and ordinary kriging was tentatively selected as the appropriate technique.

c. Natural log transformations are routinely needed for concentration data that vary over several orders of magnitude, which is common in areas of contaminant plumes. The data were transformed to log space and fit acceptable criteria for normality. After transformation to log space, the techniques described in Chapter 4 were used to guide the following steps for variogram construction:

(1) An exponential model was used to fit a directional variogram at an angle of 150 counterclockwise degrees to the east-west baseline. The variogram had a nugget of 1.00 log concentration squared, a sill of 3.20 log concentration squared, and a range of 1,295 m [Figure 5-8a and Table 5-1 (water quality A)].

(2) An exponential model was also fit to a directional variogram at an angle of 240 counterclockwise degrees to the east-west baseline. The variogram had a nugget of 1.00 log concentration squared, a sill of 3.20 log concentration squared, and a range of 229 m [Figure 5-8b and Table 5-1 (water quality A)].

(3) Cross-validation was performed using the geometric anisotropy of the two variograms and the results [Figures 5-8c and 5-8d, and Table 5-1 (water quality A)] were acceptable.

d. The residuals are symmetrically distributed, (Figure 5-8d). However, the scatterplot (Figure 5-8c) indicates that small concentrations are overestimated and that large concentrations are underestimated. This discrepancy in the estimates does not indicate an error in the model, but rather, indicates a consequence of data that have a large nugget compared to the sill; in this example the nugget is approximately 30 percent of the sill. The large nugget decreases the predictive capacity of the model and increases the smoothing introduced by kriging.

e. The established variogram then was used, along with the measured locations, to produce ordinary kriging estimates for all points in a 40-by-20 grid using a grid spacing of about 91-by-91 m. For the kriging procedure, a search radius of about 1,524 m with maximum of 16 and a minimum of 8 locations was specified. Gray-scale maps of kriging estimates, back transformed to concentrations and in log space, as well as the kriging standard deviations in log space, are shown in Figures 5-9a, 5-9b, and 5-9c.

f. The back-transformation procedure was a simple exponentiation of the log space kriging estimates. Such a back-transformation does not use bias-correction factors to deal with moment bias and, consequently, the back-transformed values must be interpreted as a median value rather than a mean value. The simple back-transformation, however, is convenient and was performed, principally, to enhance visual interpretation of the kriging estimates. Univariate statistics for the log-space kriging estimates are listed in Table 5-2 (water quality A). The kriging results do have noticeable smoothing; however, they also indicate a plume emanating from an area just northwest of the center of the area and movement, as well as some dispersion, to the southeast; the estimates are a very good representation of the results from many other more elaborate studies.

g. An additional comment concerning log transformations is appropriate. To indicate the effect of the log transform on probabilities in

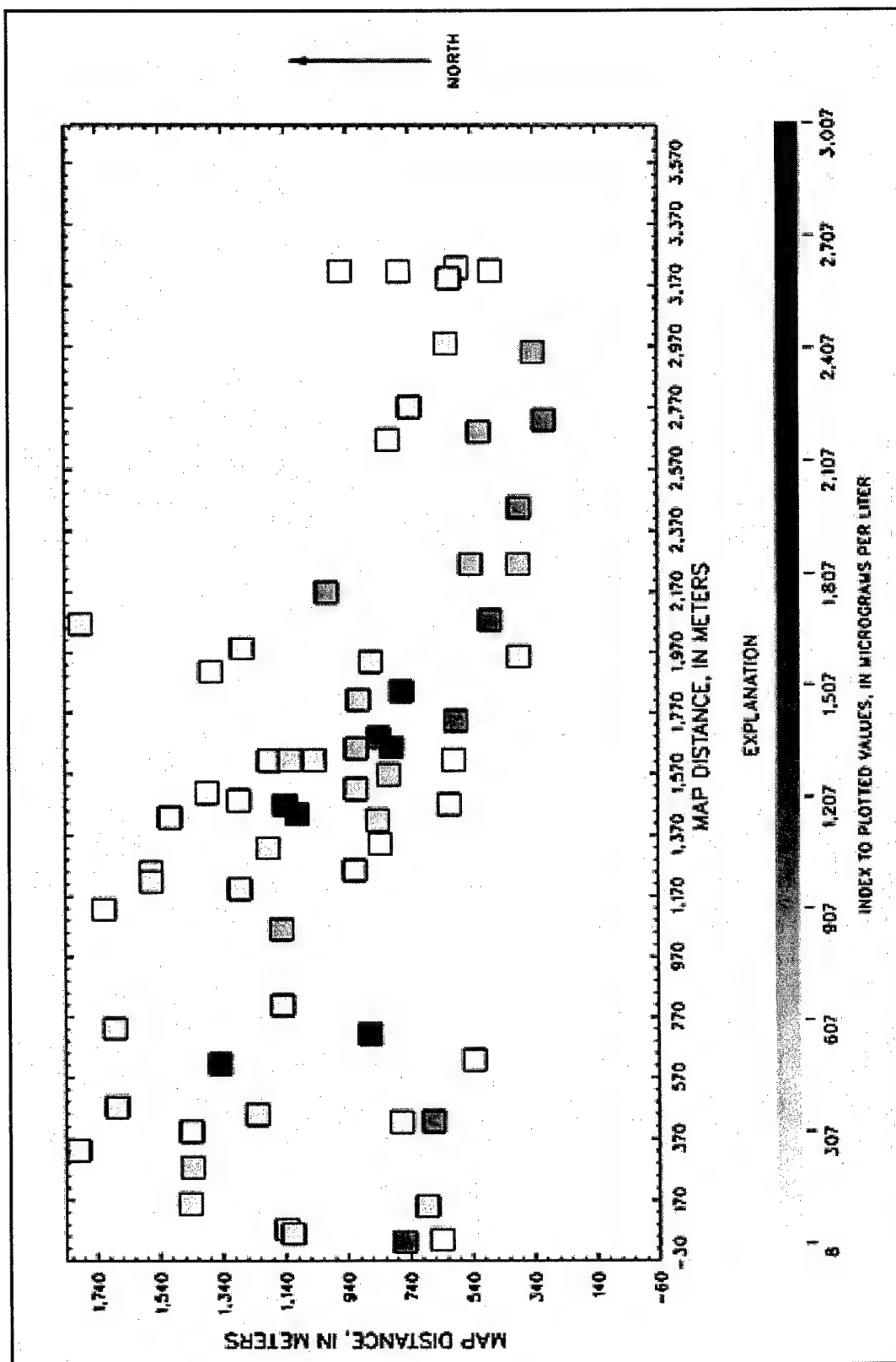


Figure 5-7. Measured data for water-quality examples

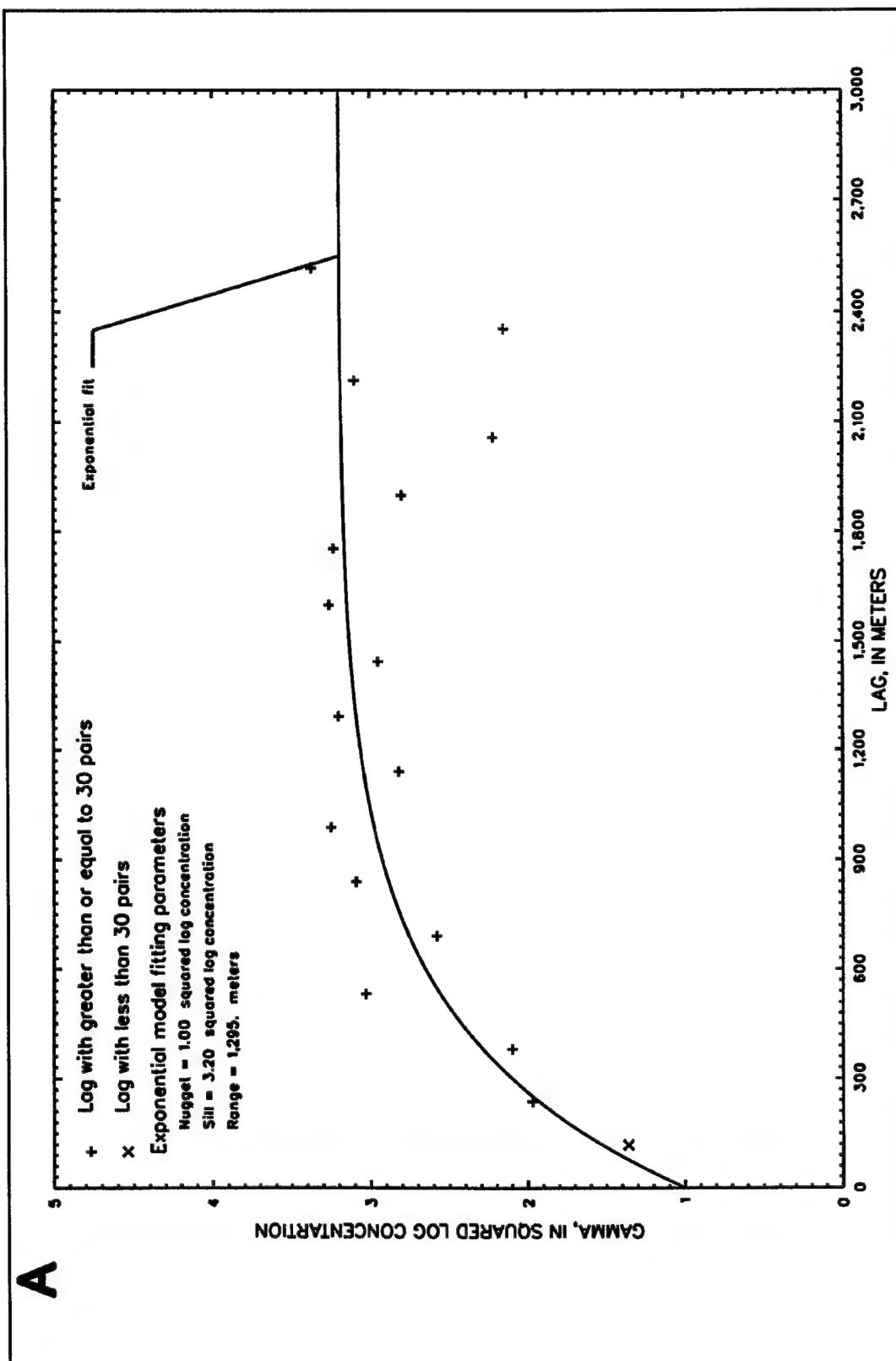


Figure 5-8. Directional variograms and variogram cross-validation plots for water-quality example--A, theoretical major-direction variogram (southeast); B, theoretical minor-direction variogram (northeast); C, cross-validation scatterplot; D, cross-validation probability plot (Sheet 1 of 4)

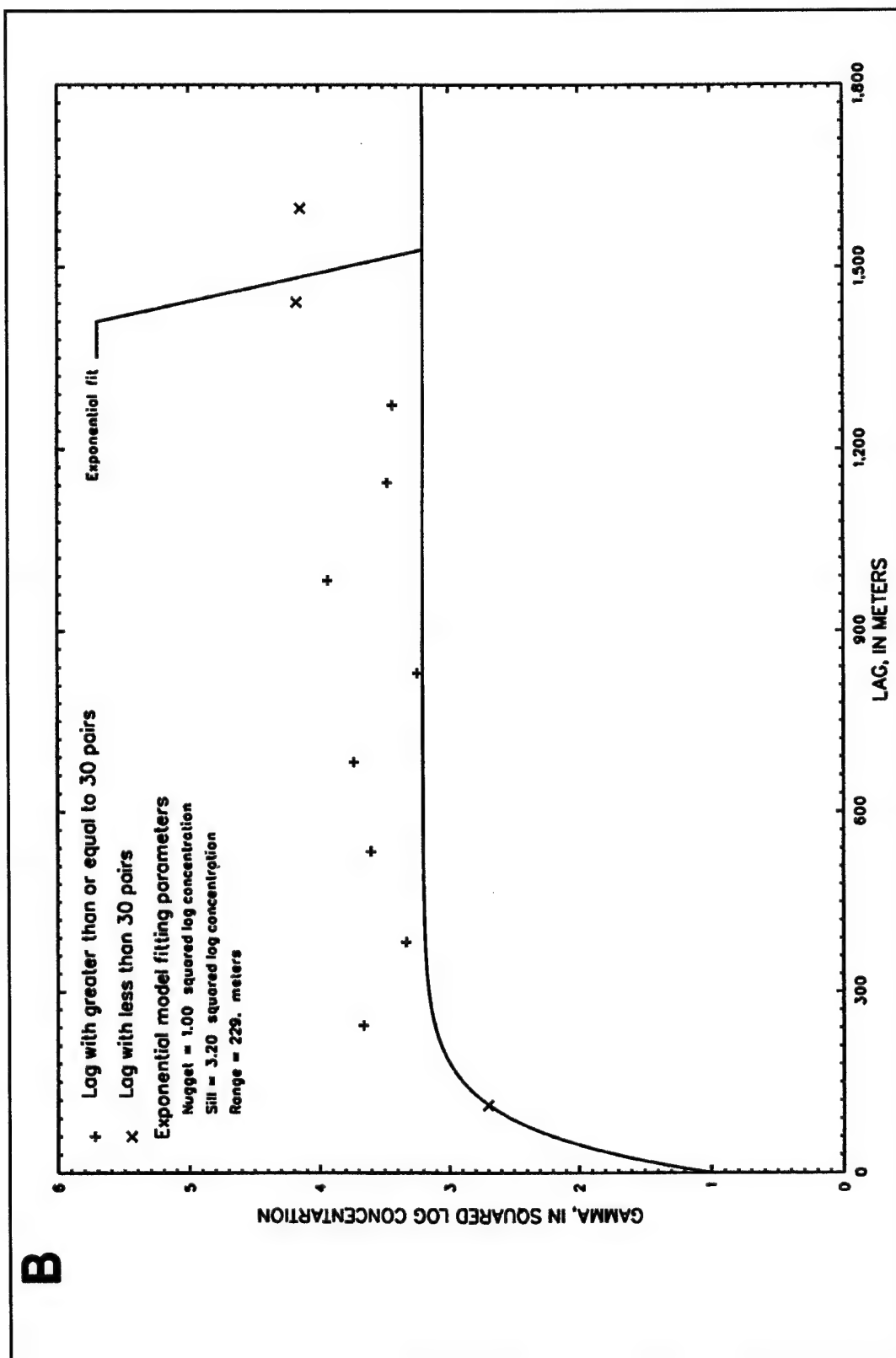


Figure 5-8. (Sheet 2 of 4)

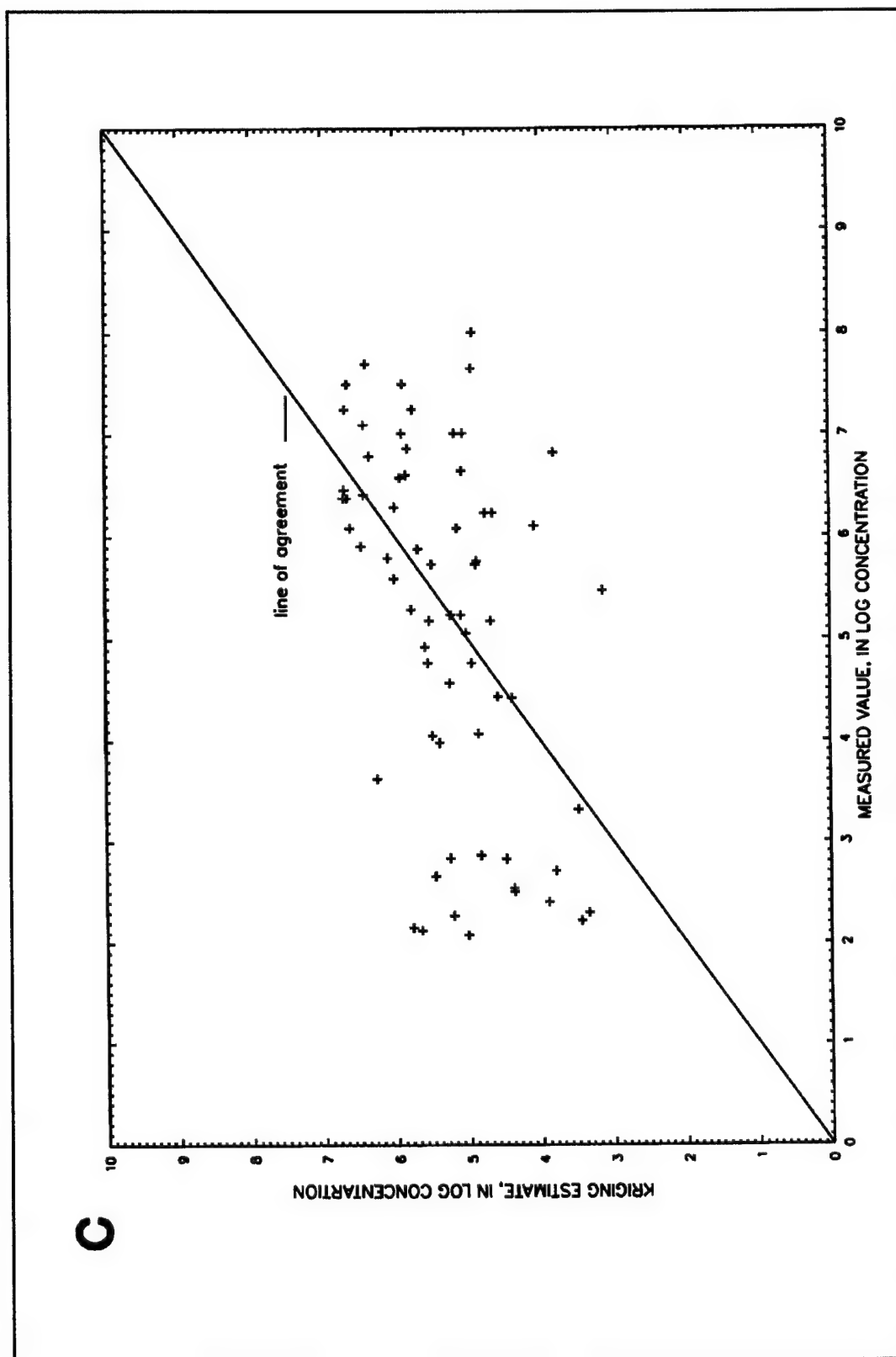


Figure 5-8. (Sheet 3 of 4)

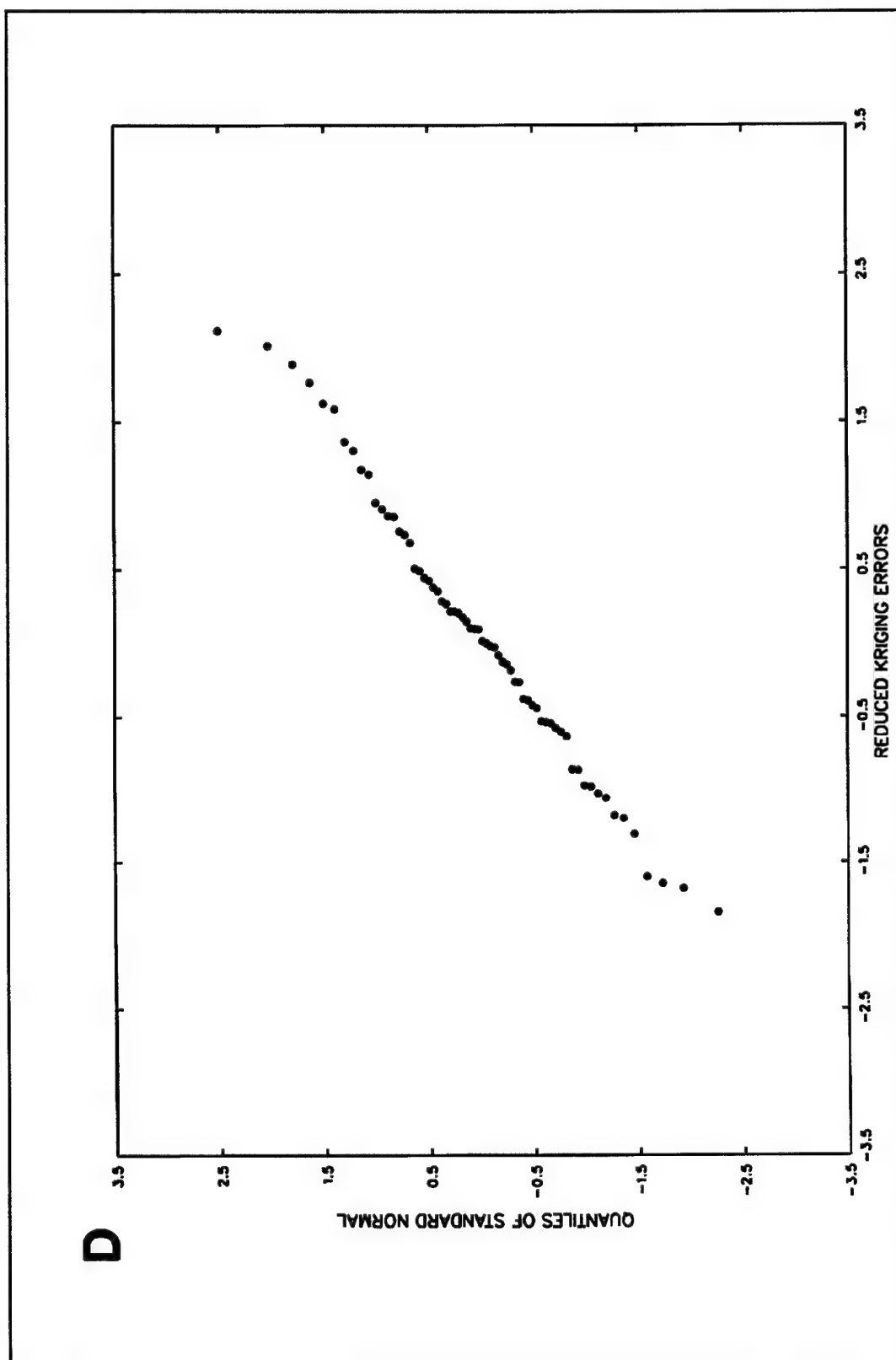


Figure 5-8. (Sheet 4 of 4)

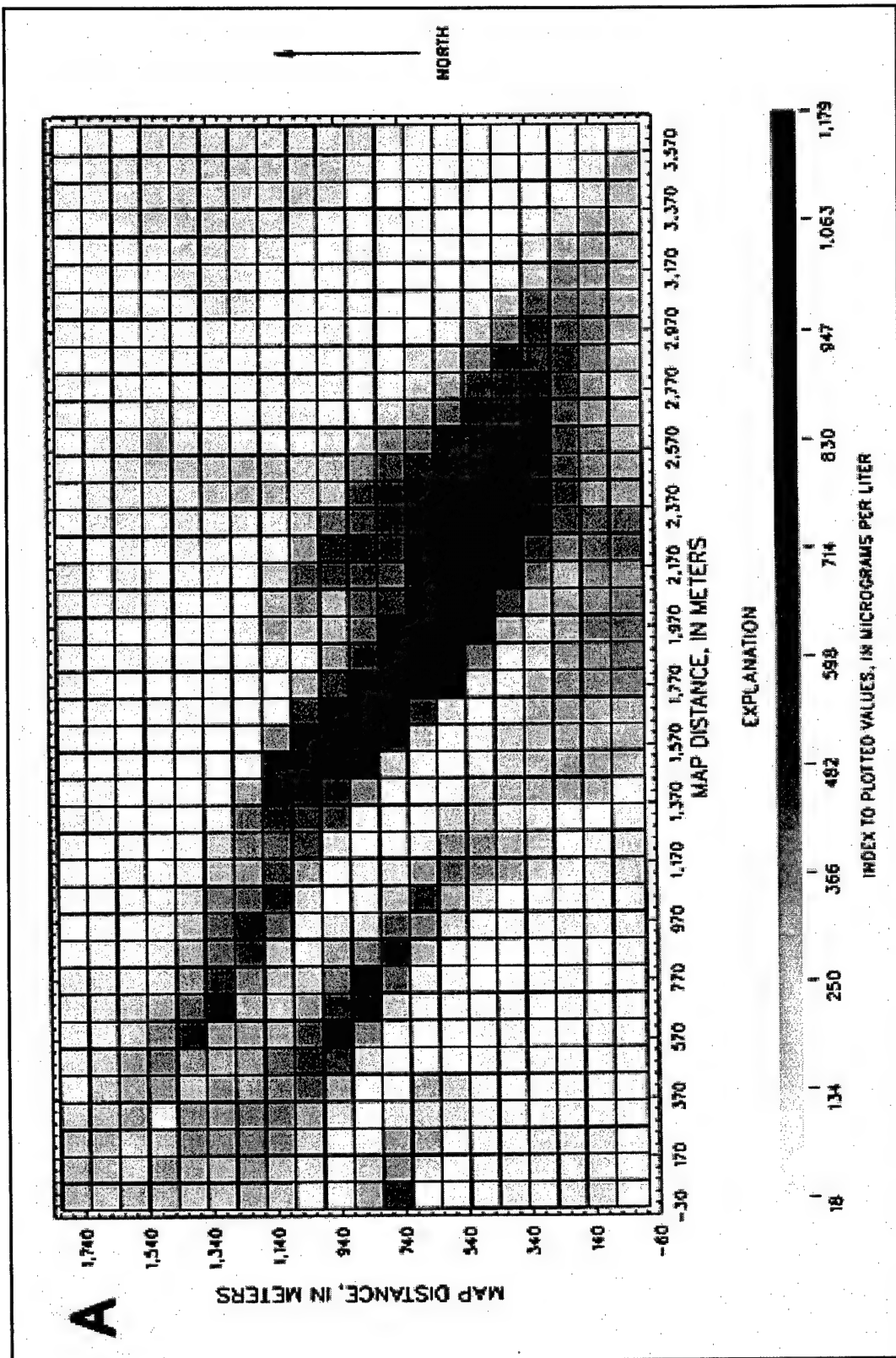


Figure 5-9. Kriging results for water-quality examples--A, kriging estimates back-transformed; B, kriging estimates in lot space; C, kriging standard deviations in lot space; D, 95-percent confidence level of kriging estimates back-transformed (Sheet 1 of 4)

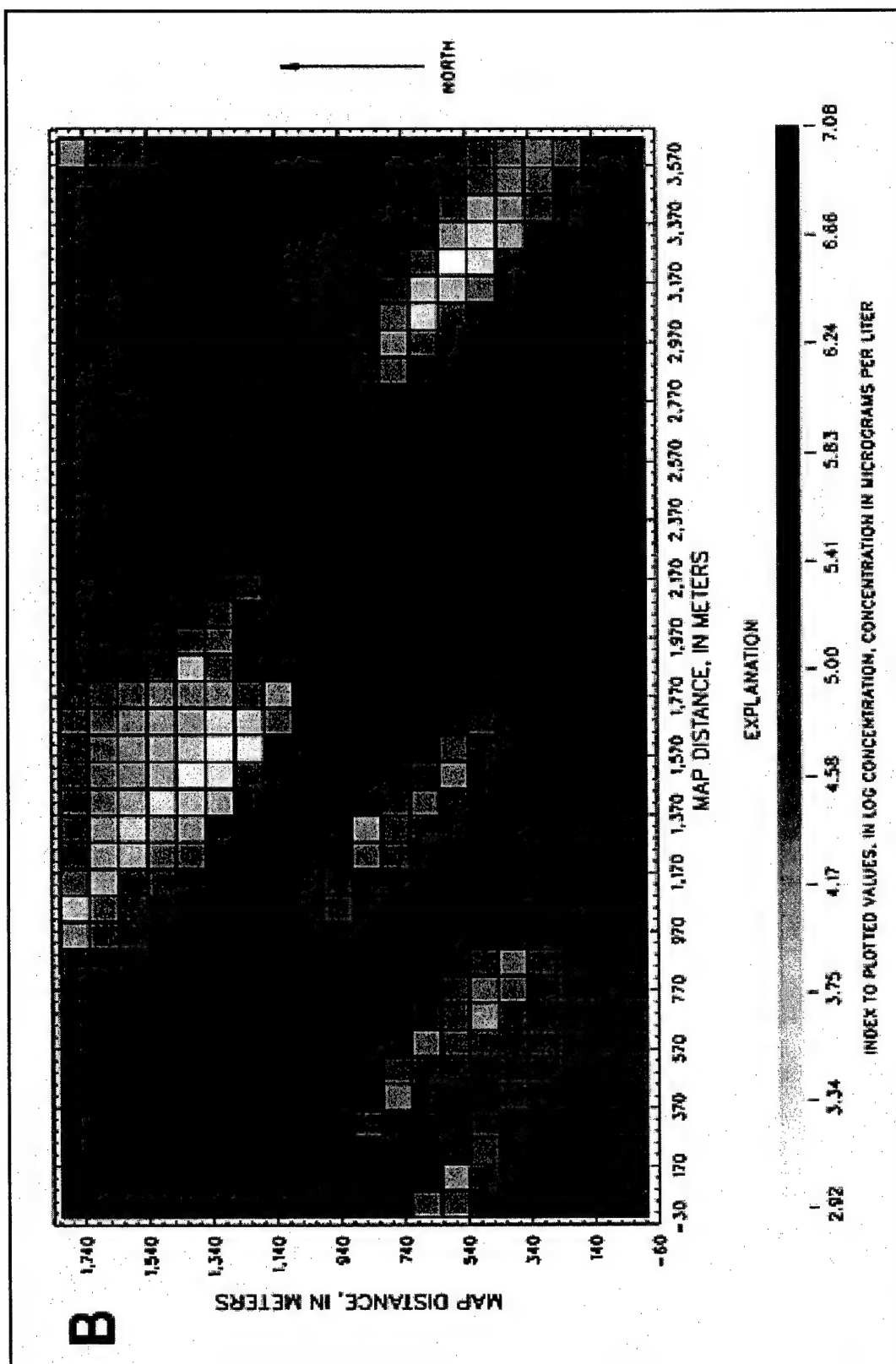


Figure 5-9. (Sheet 2 of 4)

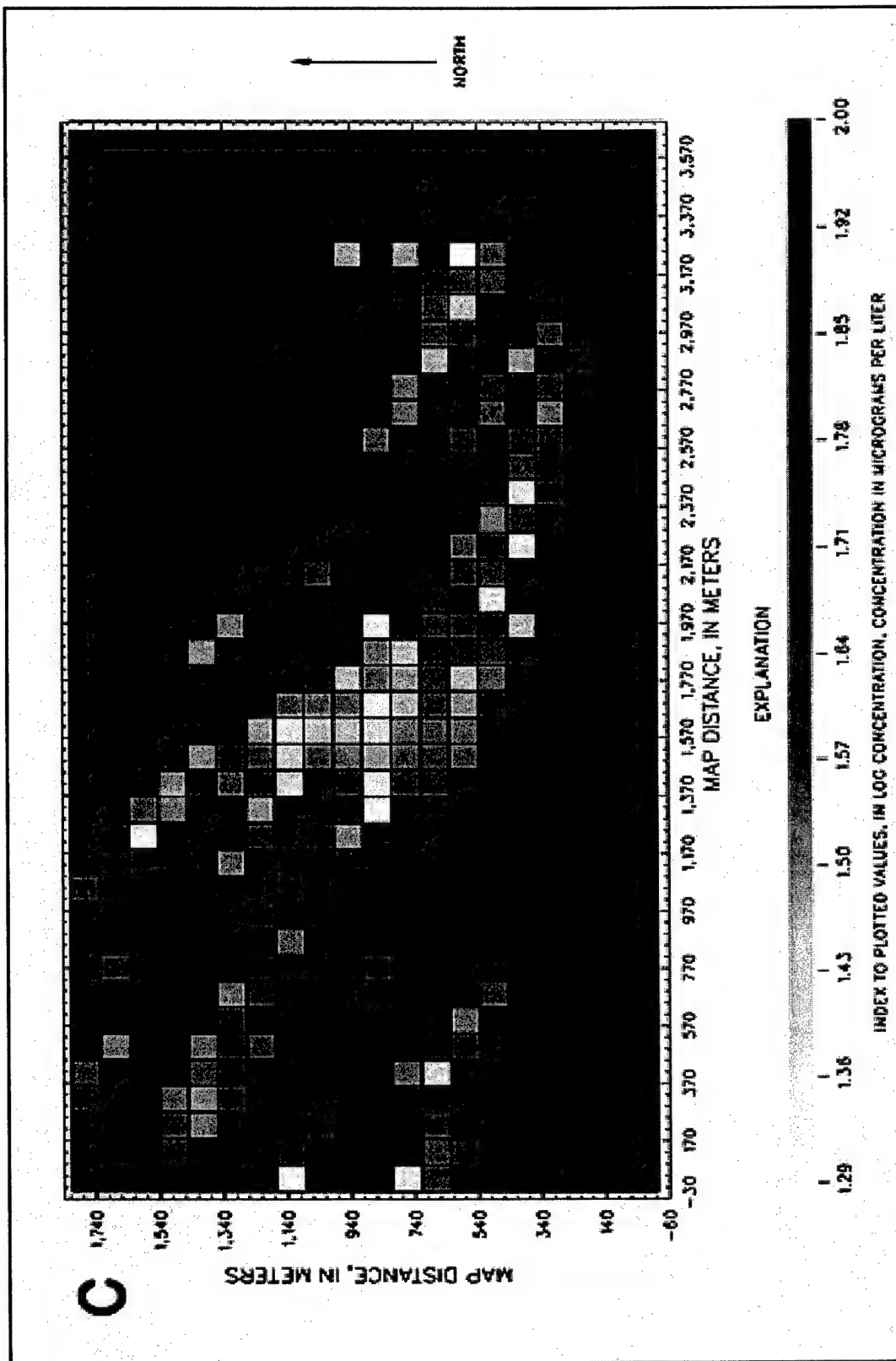


Figure 5-9. (Sheet 3 of 4)

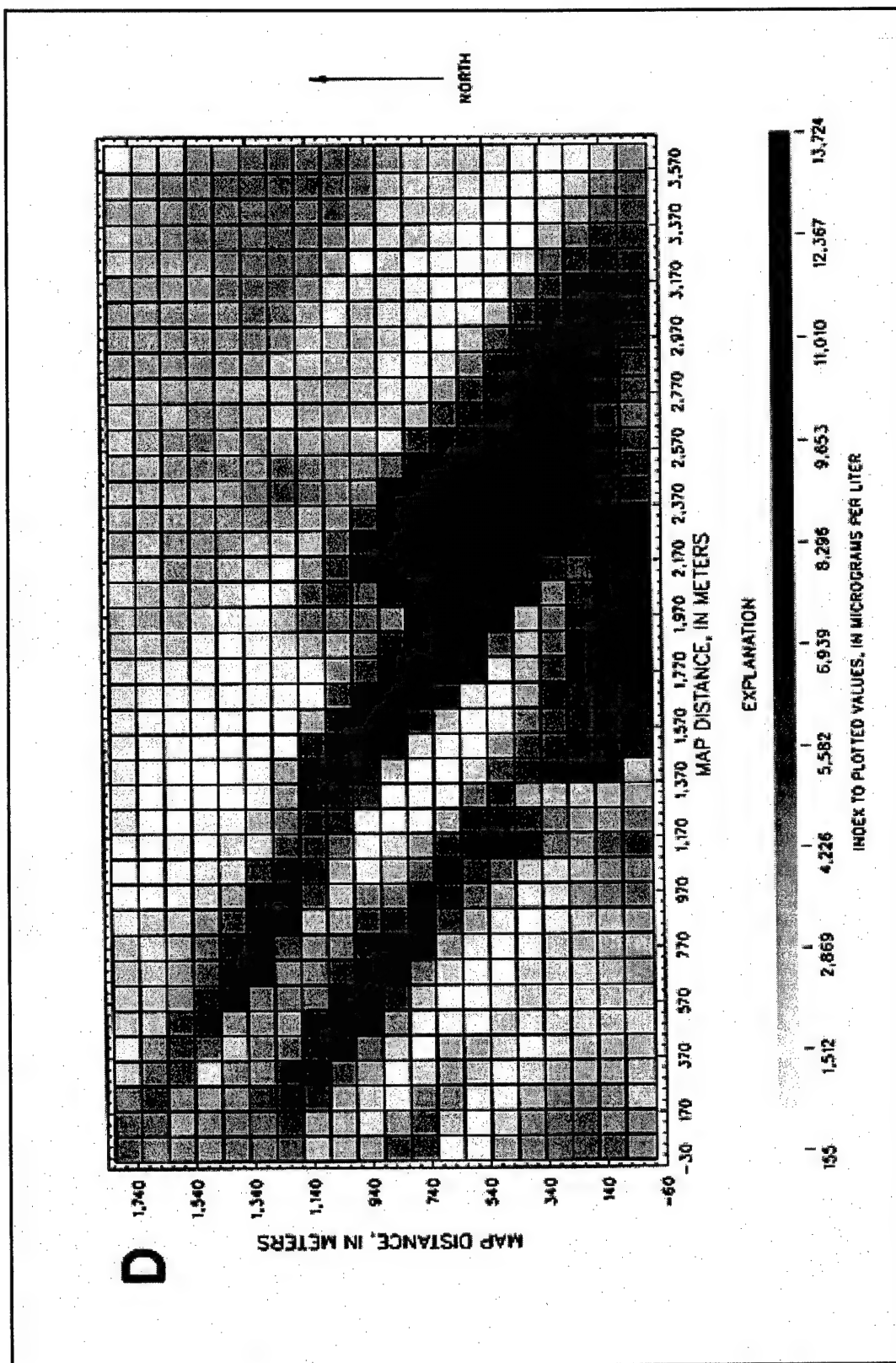


Figure 5-9. (Sheet 4 of 4)

converting, or back-transforming, kriging estimates, the kriging estimates and the kriging standard deviations, in log space, were used to estimate the one-sided 95th percentile at each kriging-estimate location according to the formula:

$$\hat{C}_{0.95} = \exp \left[\hat{Z}(x_0) + 1.645\sigma_K(x_0) \right] \quad (5-1)$$

where $\hat{Z}(x_0)$ is the kriging estimate at location x_0 , in log space, and $\sigma_K(x_0)$ is the corresponding kriging standard deviation in log space. The resulting map is shown in Figure 5-9d. Such a map can be used to indicate areas where the true concentration has only a 5-percent chance of exceeding the value shown.

h. To perform indicator kriging, the indicator transformation, as described in Chapter 2, was applied. An indicator cutoff equal to the median value of 270 for the untransformed measured data was selected. The model for indicator kriging estimates the probability that the concentration would be less than the indicator cutoff. The techniques described in Chapter 4 were used to guide the following steps in variogram construction:

(1) No trends were indicated during preliminary exploration, and ordinary kriging was tentatively selected as the appropriate technique.

(2) A spherical model was used to fit an anisotropic variogram at an angle of 150 deg counterclockwise to the east-west baseline. The variogram had a nugget of 0.05 indicator units squared, a sill of 0.25 indicator units squared, and a range of 610 m [Figure 5-10a and Table 5-1 (water quality B)].

(3) A spherical model also was fit to an anisotropic variogram at an angle of 240 deg counterclockwise to the east-west baseline. The variogram had a nugget of 0.05 indicator units squared, a sill of 0.25 indicator units squared, and a range of 213 m [Figure 5-10b and Table 5-1 (water quality B)].

i. The established variogram, along with the indicator transform of the measured data, was used to produce ordinary kriging estimates for the same grid and search criteria as the first water-quality example. A gray-scale map of the kriging estimates is shown in Figure 5-11. The kriging indicator map provides a gridded estimate for the probability of contaminant values being less than the indicator cutoff, which is a concentration of 270 in this example.

j. The cutoff value selected for the preceding indicator kriging example is probably higher than many investigators involved in HTRW site investigations would like to use. In this case the number of measurements [66 in Table 4-1 (water quality B)] used in this example, which is probably a high number of measurements for typical HTRW site investigations, would not permit construction of an indicator variogram for indicator values much lower than the median. An alternative to this problem would be to assume that the log-transformed kriging model developed in the first water-quality example is correct and to rely on the kriging estimates from that model to determine areas greater than or less than some indicator value. The same estimates also could be used to compute the probability that the concentration was less than some arbitrarily selected value.

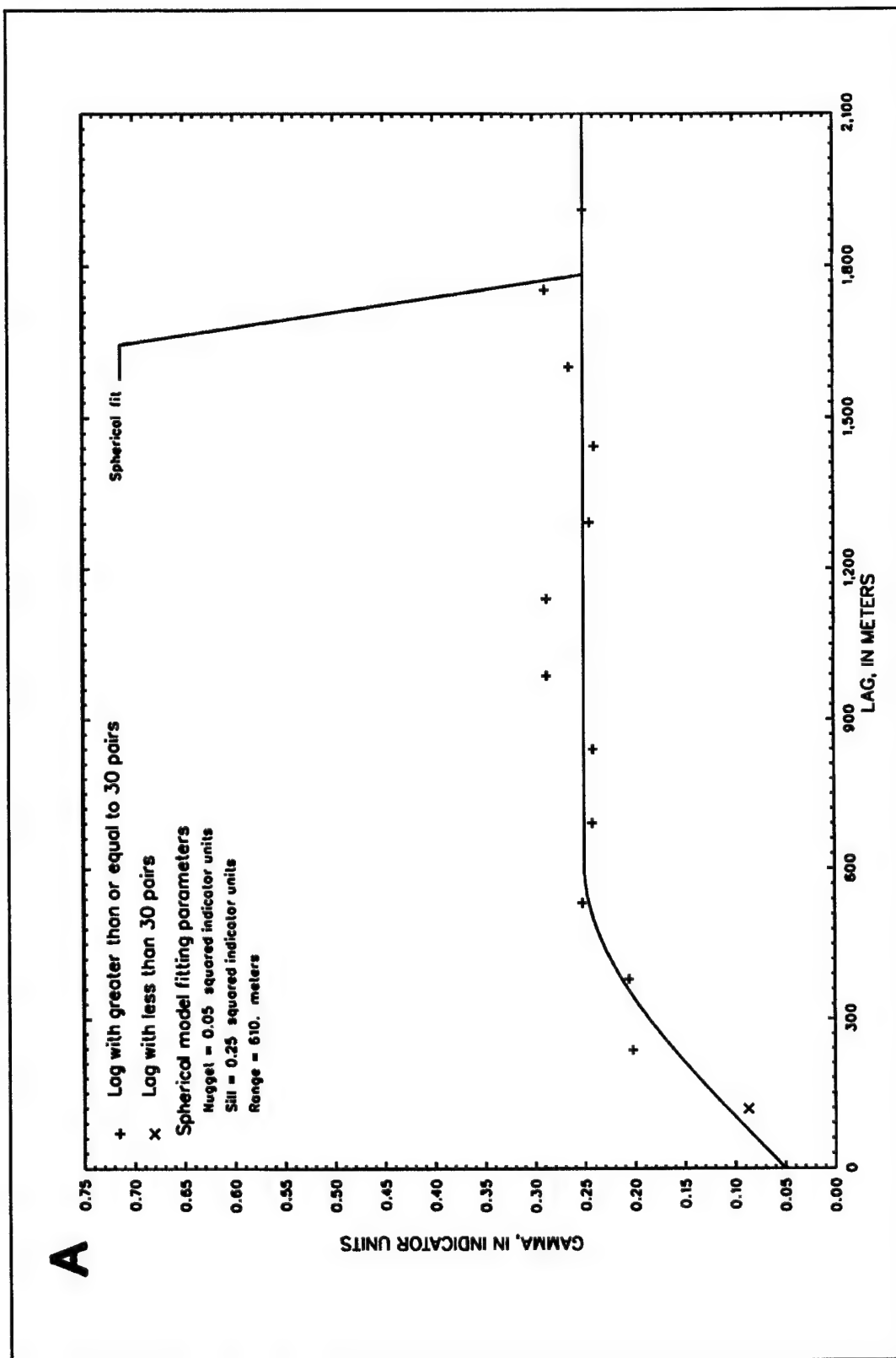


Figure 5-10. Directional variogram plots for indicator kriging water-quality example--A, theoretical major-direction variogram, and B, theoretical minor-direction variogram (Continued)

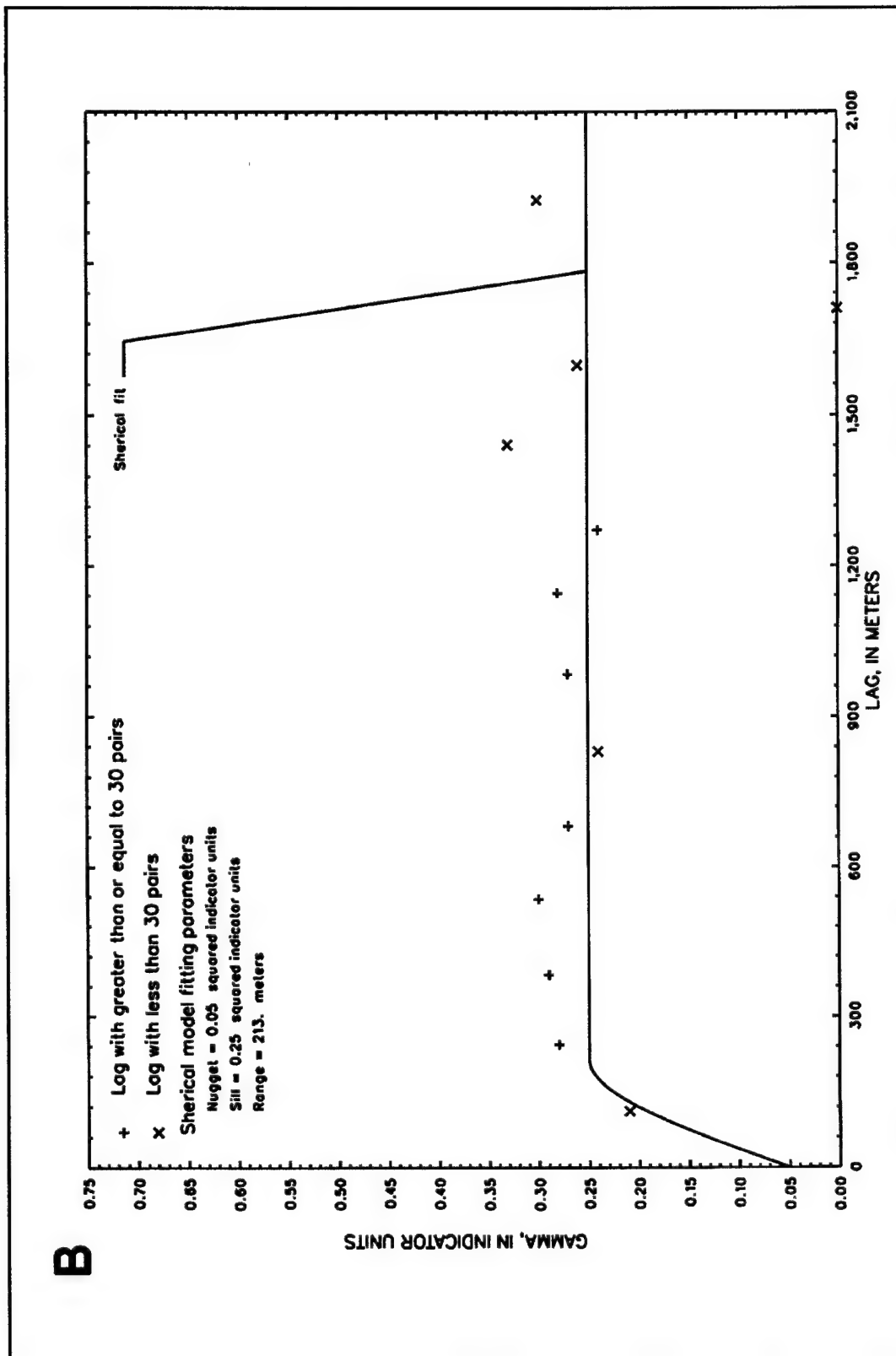


Figure 5-10. (Concluded)

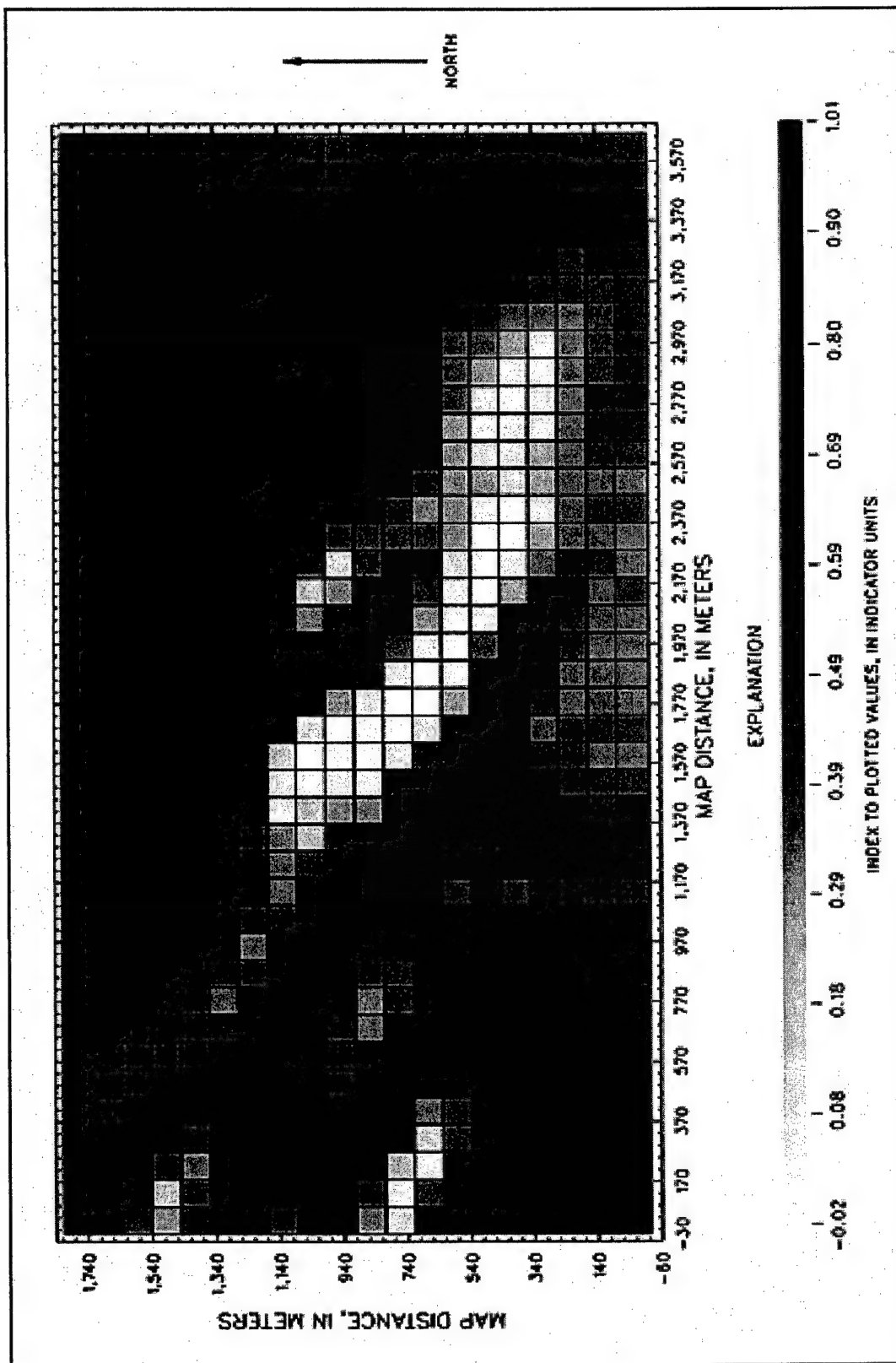


Figure 5-11. Indicator kriging results for water-quality example

Chapter 6 Review of Kriging Applications

This chapter will briefly treat three principal topics; applicability of kriging techniques, important elements that need to be addressed in kriging applications, and errors in measured data. Much of the information presented in this section has been gathered from other sections of this ETL and is presented collectively here. The items identified as important to kriging applications may be helpful in assessing kriging applications under review.

6-1. Applicability of Kriging

a. In the preceding sections of this ETL, the theory of kriging techniques has been summarized, and examples have been given to indicate the utility of kriging techniques in HTRW site investigations. The examples presented were selected so that kriging would provide satisfactory results or be applicable. Additionally, the examples were designed so that, for the purposes of demonstration, some sort of adjustment of the data was needed; that is, drift was removed or transformations were made.

b. Investigators are very likely to have data for which, although, in a strict sense, kriging may be applicable, results may be unsatisfactory. A good deal of fundamental information that may be used to establish how satisfactory application of kriging techniques might be has been presented in the preceding sections of this ETL. In particular, Chapter 4 includes a detailed discussion on variogram construction, the preliminary step in any kriging application, and systematically describes many decisions in this process that need attention. If the investigator cannot construct or otherwise obtain a variogram that has structure, then the results of a kriging application may not be satisfactory. Some additional discussion designed to guide investigators in evaluating the amount of data that may be required for kriging applications is presented in this section. This discussion will assume that the measured data are correct; a

separate and brief discussion of measurement errors will also be presented in this section.

c. Many investigators will have a tendency to focus on the amount of measured data that is available as an initial consideration. It is important for the investigator to realize that decisions concerning the applicability of kriging techniques cannot be based simply on the amount of measured data. However, unless the investigator is presented with a reliable variogram, the amount and spatial distribution of measured data can be a constraint. If, for instance, the investigator has fewer than 25 measured values at optimal locations from the field, there may not be enough data to confidently estimate Gaussian variogram parameters, a smaller amount of measured data may be suitable for other variogram models.

d. The amount of data needed to apply kriging techniques is not easy to determine, but information in this ETL, especially in Chapter 4, and the literature cited can provide some guidance. Section 4-4 points out that a good minimum for the number of pairs of locations in each variogram lag is 30 and the American Society for Testing and Materials (Standard D 5922-96) has suggested that 20 may work well also. Most investigators would probably feel comfortable defining a Gaussian form (which, because it has more inflection, is more difficult to fit compared to the other standard variogram models) with 8 to 10 optimally located sample variogram points (enough points to define the nugget, two areas of curvature, and the sill). In this ideal case, about 25 measured values would be needed to fulfill the conservative minimum of 30 pairs per lag. In this case, the relatively few measured data points need to be systematically located so that the optimally located variogram points can be computed. If the measured data were not located systematically, as is usually the case, then more measured data would be needed.

e. Once sample variogram points meeting the required number of pairs of locations can be defined, the investigator needs to have a resulting variogram that has structure. The variogram, for instance, may simply exhibit noise about a

horizontal line and have no structure. If measured data are clustered and the lags have been minimized to meet the required number of pairs of locations, the variogram may seem horizontal because it is dominated by small-scale effects in the clustered data. The investigator then has latitude to adjust the lags and attempt to balance the lag spacing and required number of pairs of locations, as described in section 4-4. However, the variogram could also seem horizontal because the actual sill is reached within a very small lag. If that lag is smaller than the minimum spacing of measured data, obtaining structure in the variogram would not be possible. If the investigator has a variogram with no structure, the measured data need to be considered independent, and kriging techniques, at the lag of the measured data, would be ineffective or at least offer little advantage over other interpolation techniques.

6-2. Important Elements of Kriging Applications

a. Many important elements of kriging applications have been discussed in this ETL. These discussions have been presented as a systematic and sequential method designed to provide guidance in kriging applications. Occasionally, an investigator will be presented with the results of a previous kriging application and will need to evaluate the application before deciding whether or not to use the results. This section presents a brief review of some important elements of kriging applications that such an investigator may use in that evaluation. For a more detailed discussion of important elements of geostatistical applications, the reader is referred to the American Society of Testing and Materials (Standard D 5549-94) for content of geostatistical investigations.

b. The presence of or lack of stationarity in the spatial mean needs to be demonstrated definitively. If the spatial mean is nonstationary, then drift is indicated and appropriate measures to establish stationarity, which are similar to the measures presented in section 4-3, need to be part of the application. In ideal situations,

nonstationarity occurs as a gradual change. HTRW site investigations may present cases, especially when dealing with water-quality data in and around plumes, that have abrupt step-like changes at plume boundaries and do not appear as regional drift. In these cases the investigator needs to be aware that without knowledge of the plume boundaries, points from within the plume will be grouped with points from outside the plume in computing the sample variogram. The effect of this problem is minimized as long as the investigator can define lags that allow data points within the plume to be grouped together.

c. The construction of the variogram needs to be described. The description needs to address the number of pairs of locations in each variogram lag and to demonstrate that the variogram has structure. A plot of the variogram is helpful to demonstrate the presence or absence of structure. The variogram construction discussion also needs to establish the presence of or lack of isotropy. If anisotropy is present, its nature needs to be established, and it needs to be addressed by variogram adjustments similar to the adjustments presented in section 4-5*b*.

d. The variogram cross-validation statistics described in section 4-9 are useful and, if available, they can aid in the evaluation of a kriging application; authoritative and definitive kriging applications should include cross-validation. Cross-validation statistics need to conform to the guidelines discussed in section 4-9. Section 4-9*b* indicates that the cross-validation exercise needs to balance minimizing the kriging cross-validation errors with efforts to guard against bias. Also, as discussed in section 4-9*b*, if probabilistic statements are part of the kriging application, there needs to be some demonstration about the normality of the reduced kriging error such as the cross-validation probability plots included with the examples in Chapter 5.

e. Maps of the kriging estimates and standard deviations need to be presented or discussed. The maps of kriging estimates need to conform to any qualitative information about the information

portrayed on the maps that is available to the investigator. The maps of kriged standard deviations can be used to determine where there are large areas of uncertainty in the kriging estimates.

f. Finally, the variogram and kriging algorithms are most useful as interpolation rather than extrapolation tools. Once the application extends to areas beyond the geographic extremes of the measured data, or perhaps those extremes plus the range, there needs to be some qualification of the area of extrapolation. For instance, in universal kriging, the practitioner would need to have some assurance that the conditions of drift defined in the study area continue into the area of extrapolation.

6-3. Errors in Measured Data

a. Data associated with HTRW site investigations have the same opportunities for errors that most investigations do. The errors may involve, among others, bias, inaccuracy, or lack of representativeness. The classical nature of these errors is described in EM 200-1-2, "Technical Project Planning," (U.S. Army Corps of Engineers 1995), which describes HTRW data-quality design.

b. The presence of contamination may complicate the function of errors in HTRW site investigations. Because these investigations often

concern contamination, there can be large ranges of values for data involving contaminant concentrations, and these large ranges have a tendency to increase the incidence of data that may seem to be statistical outliers. Even more complicating is the presence of high concentrations of organic materials that may create challenging analytical problems in laboratory determinations that also may lead to reported values that seem to be statistical outliers. In either case, the kriging practitioner is likely to find that the apparent outliers have a strong effect on the results of the kriging application.

c. When HTRW site investigations find data that seem to be outliers, the data need to be very carefully evaluated before removal is seriously contemplated. Automated outlier detection tools, as suggested in section 4-8, may best be used to identify points that may be outliers and warrant further investigation. Often data that appear to be outliers may be the most important and meaningful data of all measurements. For example, in the first case described in the preceding paragraph, apparent outliers often are representative values. In the second case, the reported value may be an erroneous determination that has been affected by the extremely contaminated nature of the sample matrix. The investigator needs to either possess or have access to qualitative or institutional knowledge that will aid in outlier interpretation.

Chapter 7 Other Spatial Prediction Techniques

7-1. General

a. In this chapter, some alternative approaches to spatial prediction are discussed. At the beginning of Chapter 2, the distinction between stochastic and nonstochastic techniques for spatial prediction was discussed. Kriging, the main subject of this ETL, is a stochastic technique because of the structure that is imposed in terms of an underlying random process (the regionalized variables) with joint probability distributions that obey certain assumptions. Kriging yields the predictor that is statistically optimal in the sense that it is the best linear unbiased predictor, given certain assumptions that are detailed in Chapter 2. There are other stochastic techniques that are less well-known than kriging in applications, such as Markov-random-field prediction and Bayesian nonparametric smoothing (see Cressie (1991)), but these will not be discussed here.

b. Several techniques that are often applied in a nonstochastic setting will be discussed. Techniques applied in such a setting are typically applied strictly empirically and not evaluated with respect to rigorous statistical criteria such as mean squared prediction error, although, as discussed in Chapter 2, such criteria may be applied in certain of the techniques such as simple average and trend analysis. It has been shown in this ETL that there are some compelling advantages for assuming some kind of stochastic setting. However, the simplicity of not having to postulate and justify the structure and assumptions inherent in stochastic analyses might be considered one advantage of nonstochastic techniques, and such an analysis may be perfectly adequate for certain problems. In addition to statistical optimality and simplicity, there are other considerations in selecting a spatial prediction technique, such as ease of computation, sensitivity to data errors, and whether the predictors are exact interpolators; that is, match the measurements exactly at the measurement locations x_1, x_2, \dots, x_n . The last property is one that needs to be

given careful consideration by the practitioner. Kriging, as it is usually applied, is an exact interpolator. Questions may be raised, however, about whether this is a desirable property if it is known that the measurements are contaminated with a considerable amount of measurement error. One advantage of stochastic methods in general is that existence of measurement error may be incorporated objectively, and, in fact, some kriging software packages (including STATPAC) have this feature, resulting in a surface that is not an exact interpolator. Several of the nonstochastic methods discussed in this section depend on a parameter that controls the deviation from exact interpolation. The ability to adjust such a parameter when using these techniques lends a degree of flexibility to the practitioner, but selecting the best value may not be straightforward and may involve considerable subjectivity on the part of the practitioner.

c. In most of the following techniques, the predictor of the process at location x_0 takes the form of a linear combination of the measurements at locations $x_i, i=1, 2, \dots, n$. Using $\tilde{Z}(x_0)$ to denote an arbitrary predictor (the notation distinguishes the predictors to be discussed in this section from the kriging predictor, which is denoted by $\hat{Z}(x_0)$), the definition of $\tilde{Z}(x_0)$ is

$$\tilde{Z}(x_0) = \sum_{i=1}^n w_i Z(x_i) \quad (7-1)$$

Although this form is the same form that is taken by the kriging predictor, the difference is in the way the coefficients w_i are computed.

7-2. Global Measure of Central Tendency (Simple Averaging)

a. The predictor for the process at any location x_0 is the simple average of the measurements; that is, the weights w_i are all equal and are given by Cressie (1991)

$$w_i = \frac{1}{n} \quad (7-2)$$

This predictor represents the smoothest possible predictor surface. In using this predictor, a certain degree of spatial homogeneity is assumed. No attempt is made to incorporate any detectable patterns (or trends) in the mean or variance of the data as a function of location, and the fact that measurements made at points that are close to each other may be related is disregarded. Such a predictor has the advantage of being very simple to compute; it needs no estimation of a variogram or other model parameters. The disadvantage is that representing the spatial field by a single value ignores much of the relevant and interesting structure that may be very helpful in improving predictions.

b. As discussed in section 2-4, if applied in a stochastic setting, this predictor would be optimal (best linear unbiased) if there is no drift and if residuals are uncorrelated and have a common variance.

7-3. Simple Moving Average

a. Let h_{i0} be the distance of x_0 from x_i , let $h_{[k0]}$ be the ordered (from smallest to largest) distances, and fix $1 \leq k \leq n$. Then the weights w_i are (Cressie 1991)

$$w_i = \begin{cases} \frac{1}{k}, & h_{i0} \leq h_{[k0]} \\ 0, & h_{i0} > h_{[k0]} \end{cases} \quad (7-3)$$

Thus, this predictor is the average of the measurements at the k nearest locations from x_0 .

b. If k is equal to n , this predictor is identical to the simple average, with weights as given in Equation 7-2. A choice of k smaller than n reflects an assumption that the predictor needs to incorporate more of the local fluctuation observed in the data, or, equivalently, that measurements at locations near x_0 should be more informative than measurements at other locations in predicting $z(x_0)$; the smaller k is, the more variable the predictor. If

$k = 1$, the predictor is an exact interpolator and is constant on the Voronoi polygons (see section 7-5) induced by the measurement locations.

c. There are several variations of this predictor. In one such variation, a distance r may be fixed (rather than fixing k) and averages over locations that are within distance r of x_0 taken. Additionally, a moving-median may be used rather than a moving average. Sorting and testing distances can slow computations relative to obtaining the simple average, and use of medians rather than means leads to a more resistant (to outliers) predictor.

7-4. Inverse-Distance Squared Weighted Average

a. The weights w_i are (Journel and Huijbregts 1978)

$$w_i = \frac{\frac{1}{h_{i0}^2}}{\sum_{j=1}^n \frac{1}{h_{j0}^2}} \quad (7-4)$$

where again h_{i0} is the distance of x_0 from x_i .

b. In the simple moving average, weights are the same, provided measurement locations are sufficiently close to the prediction location and are zero otherwise. For the inverse-distance squared method, weights are forced to decrease in a smoother manner as distance from the prediction location increases. This predictor again has the advantage of being easy to compute. Another feature of this predictor is that it is an exact interpolator. In addition, the exponent 2 of h_{i0} may be changed to any positive number, giving the user some flexibility in determining the rate of decrease of weights as a function of distance from x_0 . Isaaks and Srivastava (1989, pp. 257-259) present an example illustrating the effects on weights of changing the exponent.

7-5. Triangulation

a. To compute this predictor, the region R is partitioned into what are referred to as Voronoi polygons V_1, V_2, \dots, V_n , with V_i being the set of locations closer to measurement location x_i than to any other measurement location. If any two polygons, V_i and V_j , share a common boundary, x_i and x_j are joined with a straight line. The collection of all such lines defines what is known as the Delaunay triangulation. There will be one such triangle containing the prediction location x_0 ; the vertices of this triangle, which are measurement locations, are labelled x_j, x_k , and x_l . The spatial prediction at x_0 will be the planar interpolant through the coordinates $(x_i, z(x_i))$, $(x_k, z(x_k))$, and $(x_l, z(x_l))$. Joining x_0 and x_j, x_k , and x_l , three subtriangles are formed. The weights w_i are (Cressie 1991)

$$w_i = \begin{cases} \frac{A_i}{A_j + A_k + A_l}, & i = j, k, \text{ or } l \\ 0, & \text{otherwise} \end{cases} \quad (7-5)$$

where A_i is the area of the subtriangle opposite vertex x_i .

b. These definitions are illustrated in Figure 7-1. In this figure, the dashed lines depict the Voronoi polygons associated with points x_1, x_2, \dots, x_6 , and the solid lines define the Delaunay triangulation. Vertices of the triangle containing the prediction point x_0 are x_1, x_5 , and x_6 , and dotted lines show the subtriangles defining the associated area A_1, A_5, A_6 . For this example, j, k , and l in the general Equation 7-5 are 1, 5, and 6, so the weights assigned to points x_1, x_5 , and x_6 are, respectively,

$$\begin{aligned} w_1 &= \frac{A_1}{A_1 + A_5 + A_6}, \\ w_5 &= \frac{A_5}{A_1 + A_5 + A_6}, \text{ and} \\ w_6 &= \frac{A_6}{A_1 + A_5 + A_6} \end{aligned} \quad (7-6)$$

It is seen that the weight assigned to a point is proportional to the area of the triangle opposite the point.

c. Computation of this predictor is slower than computation of those in sections 7-2, 7-3, and 7-4. The predictor is an exact interpolator, and the surface produced is continuous, but not differentiable at the edges of the triangulation.

7-6. Splines

a. In spline modeling, the measurements are interpolated using combinations of certain so-called basis functions. These basis functions are usually taken to be piecewise polynomials of a certain degree, say k , which is determined by the user. The coefficients of these polynomials are chosen so that the function values and the first $k-1$ derivatives agree at the locations where they join. The larger k is, the smoother will be the prediction surface. Spline techniques are often applied in a non-stochastic framework; in such a context they represent a way of fitting a surface with certain smoothness properties to measurements at a set of locations with no explicit consideration of statistical optimality. There is, however, a considerable body of work in which this technique is applied in a stochastic setting. Splines may be used, for example, in nonparametric regression estimation problems (Wegman and Wright 1983).

b. A typical approach to formulating a spline problem is to pose it as an optimization problem. In one special case, it is assumed that the first two derivatives of the prediction surface exist, which is a way of imposing a certain degree of smoothness, and that the spline function minimizes

$$\frac{1}{n} \sum_{i=1}^n [z(x_i) - \tilde{z}(x_i)]^2 + \eta Q \quad (7-7)$$

where Q is a term that depends on the first two derivatives of the predictor surface. The parameter η is a nonnegative number that needs to be

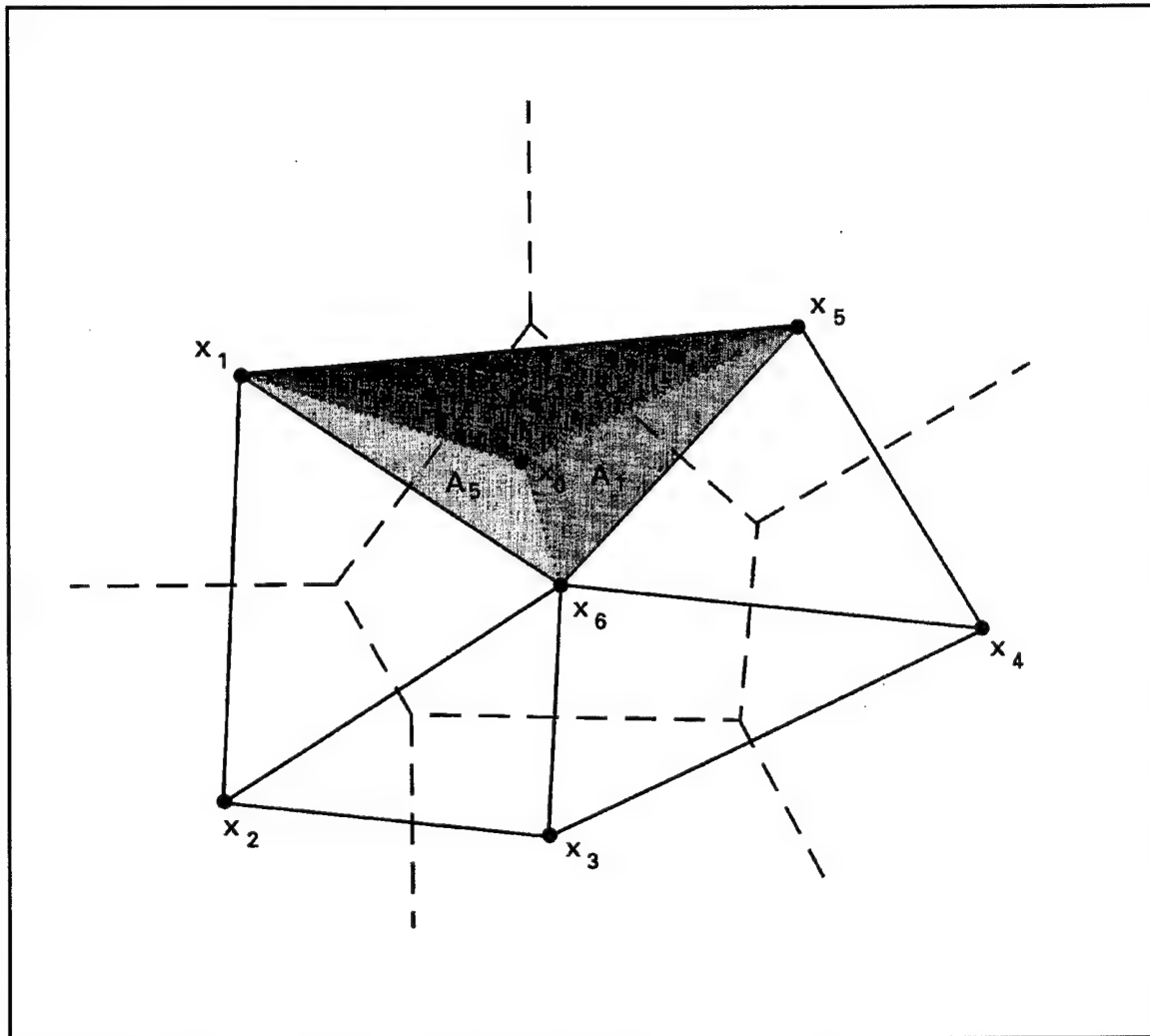


Figure 7-1. Diagram showing Voronoi polygons

specified by the user; the value of this parameter reflects the trade-off between goodness of fit to the data, measured by the first term, and smoothness, as measured by Q . If η is chosen to be 0, the spline is an exact interpolator and passes through all the data points. If $\eta > 0$, the spline is not an exact interpolator. (Splines that are not exact interpolators are referred to as smoothing splines.) There are a number of numerical procedures that may be used for fitting splines, but allowing the

smoothing parameter η to be > 0 renders the computational problem more complex.

c. Under some conditions a solution to the optimization problem (Equation 7-7) may also be obtained by a kriging algorithm if the smoothing parameter η is taken to be equal to the variance of measurement error and if a special form is chosen for the covariance function. Therefore, in this situation, spline approximation is a special type of

kriging. However, the variogram that needs to be used in the kriging equations to make the kriging predictor equivalent to the spline predictor is determined by the basis functions selected for the spline. Because the type of basis functions used is subjective on the part of the user, the resulting equivalent variogram may not be representative of the true variogram of the data. Because kriging uses the data to indicate reasonable variogram choices, kriging has an important advantage over splines. Another advantage of placing the problem in the kriging framework is the interpretation of the smoothing parameter in terms of measurement errors. In many cases, an objective estimate of the magnitude of measurement error can be obtained. The connections between kriging and splines are discussed further by Wegman and Wright (1983), Watson (1984), and Cressie (1991).

7-7. Trend-Surface Analysis

a. Trend-surface analysis is the process of fitting a function, such as that in Equation 2-43, using least squares to determine the coefficients that yield the best fit. Computationally, trend-surface analysis is equivalent to universal kriging with an assumption that the $Z^*(x_i)$ are uncorrelated. Thus, there is no need to estimate a variogram, and readily available regression packages may be used for estimating the coefficients. As in universal kriging, polynomial surfaces are the most commonly used.

b. When trend surfaces are applied in a stochastic setting, the resulting predictor will be optimal if deviations from the surface are uncorrelated and have a common variance.

7-8. Simulation

a. Consider again a regionalized random variable $Z(x)$, where x is a location in a two-dimensional study region R . Kriging is an interpolation algorithm that yields spatial predictions $\hat{Z}(x)$ that are best, or optimal, in the sense that has been discussed at some length in this ETL. The

mean-squared prediction error is smallest among all predictors that are linear in the measurements. This optimality property is local, in that the mean-squared error of predictions at unsampled locations considered one at a time is minimized, without specific regard to preservation of global spatial features. If, however, the actual realization $z(x)$ could be compared to the kriged prediction surface based on n measured values, the kriged surface would be much smoother than the actual surface, especially in regions of sparser sampling. Thus, the kriged surface will be a good and realistic representation of reality in the sense that the n measured values are honored, but it will be less realistic with respect to global properties, such as overall variability.

b. The purpose of simulation is to produce one or more spatial surfaces (realizations) that are more realistic in preserving global properties than the surface produced by interpolation algorithms, such as kriging. These realizations are produced by using numbers that are drawn randomly (Monte Carlo) to impart variability to the simulated surface, making the simulated surface more realistic in preserving the overall appearance of the actual surface. Generally speaking, simulation uses the idea that the true value of a random surface may be expressed as the sum of a predicted value (which is obtained by kriging) plus a random error, which varies spatially and depends on the random numbers drawn. Generally a number of independent realizations will be generated, and these realizations will be taken to be equally probable representations of reality.

c. A simulation algorithm is said to be conditional if the resulting realizations agree with the measurements at measurement locations x_1, x_2, \dots, x_n . If the underlying process $Z(x)$ is assumed to be Gaussian (or if a transformation may be found that makes the process Gaussian), the most common method of conditional simulation is known as sequential Gaussian simulation (Deutsch and Journel (1992), pp. 141-143). Another, more complicated, Gaussian simulation method that is particularly useful for three-dimensional simulations because of its computational efficiency is the

turning-bands method (Deutsch and Journal 1992, Journal and Huijbregts 1978).

d. In sequential Gaussian simulation a set of grid points for which simulated values are desired is defined and the points are addressed sequentially from location to location along a predetermined path. At each location, a specified set of neighboring conditioning data is retained, including the original data and simulated grid-location values at previously traversed grid locations along the path. Then, a random number is generated from a Gaussian distribution with conditional mean and variance determined using a kriging algorithm, and the value of the random number determines the simulated process at this location. The conditional Gaussian distribution used in simulation is identical to the conditional distribution discussed in section 2-6b. An idea of the computational requirements can be obtained from the fact that a kriging algorithm needs to be applied for each simulation location. For multiple realizations, if the path connecting the grid points is kept the same, the kriging equations need to be solved for only the first simulation. However, implementation of this procedure needs to take into consideration the assumptions concerning the existence of drift; the details of such an implementation are beyond the scope of this ETL.

e. A sequential algorithm like this may also be applied in the context of indicator kriging (see section 2-6c). At each grid point along the path, a (Bernouli) random variable taking on only two possible values, 0 or 1, is generated, with the relative probability of these two values being determined by indicator kriging applied, as in the previous paragraph, to the original observed indicator data and the previously simulated indicator values.

f. To get an idea of how simulation results might be used in a risk-assessment setting, assume again that the underlying process is Gaussian and that 1,000 conditional realizations have been generated. If a single grid point x_0 (which is not a measurement point) is considered, then the simulation has produced 1,000 values at x_0 , which,

when analyzed in histogram form, approximates the probability distribution of potential measurements at that location. If an interval with exactly 25 (2.5 percent) of the values less than the lower end and 25 of the values larger than the upper end were constructed, the interval would almost correspond, as expected, to the 95-percent prediction interval to $\hat{Z}(x_0) - 1.96\sigma_K(x_0)$ to $\hat{Z}(x_0) + 1.96\sigma_K(x_0)$ discussed in section 2-6b. Thus, for this single location, the simulation has not produced much more information than kriging alone would have produced. The real value of simulation, however, is that realizations not just at a single location, but at all of the grid locations jointly, are obtained. These realizations can be used to calculate probabilities associated with any number of spatial locations together. For example, the probability that the largest (maximum) contaminant value over a certain subregion is greater than a particular concentration might be assessed. (If the word "largest" here were replaced with "average," then block kriging could be used to obtain the answer.)

g. A central point that needs to be emphasized is that simulation is especially useful when probabilities associated with complicated, usually nonlinear, functions of the regionalized variables over a region need to be analyzed. The maximum function mentioned in the preceding paragraph is one simple example. For another example, consider the problem of determining placement of groundwater monitoring wells to detect and monitor groundwater contamination emanating from a potential point source. Given an existing set of hydraulic-head data, kriging might be applied and flow lines determined from resulting hydraulic-head gradients. Intersection of the flow line from the point source with the regional boundary then might be used to determine monitoring well placement. Conditional simulation would be useful to determine uncertainty associated with location of well placement or to give an indication of how many monitoring wells might be appropriate. In this case, the variable of interest, well location, is a complicated function of hydraulic heads so this is a problem for which simulation is well-suited. The reader may refer to Easley, Borgman, and Weber

(1991) for a more detailed discussion of this type of application.

h. The complicated functions of interest in groundwater studies often involve physically based groundwater flow models. Conditional simulation may be used, for example, to generate a suite of

hydraulic-conductivity realizations to be used as input to a model that produces as output a set of corresponding hydraulic-head realizations. Weber, Easley, and Englund (1991) discuss how groundwater modeling might be used with conditional simulation to study the monitoring-well-placement problem discussed in the preceding paragraph.

Appendix A: References

American Society for Testing and Materials

American Society for Testing and Materials.

"Standard guide for analysis of spatial variation in geostatistical site investigations, D 5922- 96,"

American Society for Testing and Materials Committee D-18 on Soil and Rock, Philadelphia, PA.

American Society for Testing and Materials

American Society for Testing and Materials.

"Standard guide for content of geostatistical site investigations, D 5549-94," American Society for Testing and Materials Committee D-18 on Soil and Rock, Philadelphia, PA.

Bras, and Rodriguez-Iturbe 1985

Bras, R. L., and Rodriguez-Iturbe, I. 1985. "Random functions and hydrology," Addison-Wesley Publishing Company, Reading, MA.

Clark 1979

Clark, I. 1979. "Practical geostatistics," Applied Science Publishers, London.

Cliff and Ord 1981

Cliff, A. D., and Ord, J. K. 1981. "Spatial processes, models and applications," Pion Limited, London.

Cressie 1991

Cressie, N. 1991. "Statistics for spatial data," Wiley, New York.

David 1977

David, M. 1977. "Developments in geostatistics; 2: Geostatistical ore reserve estimation," Elsevier, Amsterdam.

Davis 1986

Davis, J. C. 1986. "Statistics and data analysis in geology," Wiley, New York.

Delhomme 1978

Delhomme, J. P. 1978. "Kriging in the hydro-sciences: Advances in water resources," Vol. 1, 251-266.

Deutsch and Journel 1992

Deutsch, C. V., and Journel, A. G. 1992.

"GSLIB, geostatistical software library and user's guide," Oxford University Press, New York.

Devore 1987

Devore, J. L. 1987. "Probability and statistics for engineering and the sciences," Cole Publishing Company, Monterey, CA.

Easley, Borgman, and Weber 1991

Easley, D. H., Borgman, L. E., and Weber, D.

1991. "Monitoring well placement using conditional simulation of hydraulic head," *Mathematical Geology* 23(8), 1059-1080.

Englund and Sparks 1991

Englund, E. J., and Sparks, A. R. 1991. "GEO-EAS (Geostatistical Environmental Assessment Software) users guide," Environmental Monitoring Systems Laboratory Report EPA/600/8-91/008, Las Vegas, NV [Available from National Technical Information Service, Springfield, VA 22161 as NTIS report PB89-151 252].

Grundy and Miesch 1987

Grundy, W. D., and Miesch, A. T. 1987. "Brief description of STATPAC and related statistical programs for the IBM personal computer," U.S. Geological Survey Open File Report 87-411.

Hawkins 1980

Hawkins, D. M. 1980. "Identification of outliers," Chapman and Hall, London.

Isaaks and Srivastava 1989

Isaaks, E. H., and Srivastava, R. M. 1989. "An introduction to applied geostatistics," Oxford University Press, New York.

Jones and Zhang 1994

Jones, R. H., and Zhang, Y. 1994. "Spatial and temporal analysis of the Rocky Flats soil plutonium data," Department of Public Health and Environment, Denver, CO.

Journal 1993

Journal, A. G. 1993. "Geostatistics for the environmental sciences," U.S. Environmental Protection Agency, Las Vegas, NV.

Journal 1988

Journal, A. G. 1988. "Non-parametric geostatistics for risk and additional sampling assessment," *Principles of environmental sampling*, L. Keith, ed., American Chemical Society, 45-72.

Journal and Huijbregts 1978

Journal, A. G., and Huijbregts, C. 1978. "Mining geostatistics," Academic Press, London.

Krige and Magri 1982

Krige, D. G., and Magri, E. J. 1982. "Studies of the effect of outliers and data transformations on variogram estimates for a base metal and a gold ore body," *Mathematical Geology* 14(6), 557-564.

Lenfest 1986

Lenfest, L. W. 1986. "Ground-water levels and use of water for irrigation in the Saratoga valley, south-central Wyoming, 1980-8," U.S. Geological Survey Open-File Report 84-4040.

Myers, Begovich, Butz, and Kane 1980

Myers, D. E., Begovich, C. L., Butz, T. R., and Kane, V. E. 1980. "Application of kriging to the hydrogeochemical data from the National Uranium

Resource Evaluation Program," Union Carbide Corporation, Nuclear Division Oak Ridge Gaseous Diffusion Plant, Oak Ridge, TN.

Olea 1991

Olea, R. A. 1991. "Geostatistical glossary and multilingual dictionary," Oxford University Press.

Ripley 1981

Ripley, B. D. 1981. "Spatial statistics," Wiley, New York.

Ross 1987

Ross, S. M. 1987. "Introduction to probability and statistics for engineers and scientists," Wiley, New York.

U.S. Army Corps of Engineers 1995

U.S. Army Corps of Engineers. 1995. "Technical project planning--guidance for hazardous toxic and radioactive waste data quality design," U.S. Army Corps of Engineers Engineer Manual 200-1-2, Washington, DC.

Watson 1984

Watson, G. S. 1984. "Smoothing and interpolation by kriging and with splines," *Mathematical Geology* 16(6), 601-615.

Weber, Easley, and Englund 1991

Weber, D., Easley, D., and Englund, E. J. 1991. "Probability of plume interception using conditional simulation of hydraulic head and inverse modeling," *Mathematical Geology* 23(2), 219-239.

Wegman and Wright 1983

Wegman, E. J., and Wright, I. W. 1983. "Splines in statistics," *American Statistical Association Journal* 78(382), 351-365.

Appendix B: Notation

a	Angle for directional variogram	$C(\underline{x}_1, \underline{x}_2)$	Covariance of data values at locations \underline{x}_1 and \underline{x}_2
b	Slope of variogram	D_{ij}	Difference in values between data points i and j
c	Generic constant used for cutoff value in probability distribution or indicator transformation	E	Expectation
e	Kriging error	$I(\cdot)$	Indicator function
\tilde{e}	Reduced kriging error	K	Number of variogram bins
f	Explanatory variables used in drift equations	$N(\cdot)$	Number of squared differences in variogram bin
g	Nugget of variogram	P	Probability
h	Lag or distance between two data points	S_n^2	Sample variance on n observations
n	Number of data points	V	Voronoi polygon
m	Number of locations in a given block	Var	Population variance
r	Range of variogram	$W(\underline{x})$	Co-kriging random variable at location \underline{x}
s	Sill of variogram	$Y(\underline{x})$	Transformed variable at location \underline{x}
w	Weight	Z	Regionalized random variable
$\underline{x}(u, v)$	Location in terms of coordinates u and v	$Z(\underline{x})$	Potential value of Z at location \underline{x}
$z(\underline{x})$	Measurement of Z at location \underline{x}	$\hat{Z}(\underline{x})$	Predictor or estimate of Z at location \underline{x} , obtained from kriging
$\hat{z}(\underline{x}_1)$	Kriging estimate using measured data	$Z^*(\underline{x})$	Residuals of $Z(\underline{x})$
A	Area of triangle	$\hat{\hat{Z}}(\underline{x})$	Arbitrary predictor of Z at location \underline{x}
B	Area designation in block kriging	\bar{Z}_n	Sample mean of n observations
C	Population covariance function	β	Regression coefficient used in polynomial representation for drift
\hat{C}	Sample covariance function		

$\hat{\gamma}$	Sample variogram	$\sigma(\underline{x})$	Spatial standard deviation at location \underline{x}
γ	Theoretical variogram	$\sigma^2(\underline{x})$	Spatial variance at location \underline{x}
$\gamma(h)$	Theoretical variogram for lag h	$\sigma_k(\underline{x})$	Kriging standard deviation at location \underline{x}
λ	Optimization coefficient	$\sigma_k^2(\underline{x})$	Kriging variance at location \underline{x}
η	Parameter used in spline analysis	$\mu(\underline{x})$	Spatial mean at location \underline{x}
$\rho(h)$	Correlation function as function of h		

CEMP-RT

DEPARTMENT OF THE ARMY
U.S. Army Corps of Engineers
Washington, DC 20314-1000

ETL 1110-1-175

Technical Letter
No. 1110-1-175

30 June 1997

Engineering and Design
PRACTICAL ASPECTS OF APPLYING GEOSTATISTICS
AT HAZARDOUS, TOXIC, AND RADIOACTIVE WASTE SITES

Distribution Restriction Statement

Approved for public release; distribution is unlimited.